METHODS FOR STREAMLINING EXPENSIVE FUNCTION CALLS

by

Benjamin Jeppson Taylor

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Chemical Engineering

The University of Utah

August 2011

Copyright © Benjamin Jeppson Taylor 2011

All Rights Reserved

The University of Utah Graduate School

STATEMENT OF THESIS APPROVAL

The thesis of **Benjamin Jeppson Taylor**

has been approved by the following supervisory committee members:

Donald Keith Roper	, Chair	**12/23/10** Date Approved
James C. Sutherland	, Member	Date Approved
Edward M. Trujillo	, Member	**01/12/11** Date Approved

and by **JoAnn S. Lighty** , Chair of the Department of **Chemical Engineering**

and by Charles A. Wight, Dean of The Graduate School.

ABSTRACT

The focus of this thesis has been developing and validating new methods of streamlining expensive function calls. These new methods of streamlining are applied to the problem of modeling electrodynamic response in infinite two-dimensional nanoparticle (NP) arrays. Using these methods, an exact analytical solution for electromagnetic interaction with two-dimensional spherical periodic infinite arrays has been developed. The new solution offers fundamental insight into the singularities responsible for dramatic sensitivity increases seen in sensory applications. The new analytical solution also provides increased computational time with 60x improvement demonstrated in this work. In addition, a new electroless deposition method followed by thermal annealing was developed by Ahn et al. for spherical Au array fabrication. To complement the model and fabrication technique, additional improvements in NP array sample analysis have been adapted and developed. Algorithms for NP particle density, size, and shape analysis are described using Atomic Force Microscope (AFM) and Scanning Electron Microscope (SEM) with a comparison between the two microscopy methods. Using fabrication and analysis, the methods of streamlining expensive function calls can be tested and improved.

CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

The author would like to acknowledge Dr. Randy Polson at the Dixon Laser Institute of the University of Utah, Salt Lake City, Utah, USA, for technical assistance with AFM, and also Analia G. Dall' Asen and Wonmi Ahn for their close collaboration in the lab and data analysis, and Dr. Keith Roper for his support and guidance.

CHAPTER 1

INTRODUCTION

1.1 Nanoscience Introduction

As scientific research continuous to find bottlenecks in computing power, care must be taken to ensure maximum efficiency within the algorithms themselves. This work demonstrates a novel method of streamlining computationally expensive function calls to improve simulation results. This concept is demonstrated with an algorithm used for simulating electrodynamic response of two-dimensional nanoparticle (NP) arrays. Accurately simulating the electric field associated with these dipole array sensors allows for sensor design optimization and increased sensitivity. By increasing the sensor sensitivity, antigen detection on the order of a single molecule may be possible. In order to validate the new results, NP array fabrication as well as metrology techniques were used. For metrology, this work covered UV-VIS spectroscopy, Scanning Electron Microscopy (SEM), and Atomic Force Microscopy (AFM) as well as new image processing algorithms that were needed to generate summary statistics. The new analytical solution that was derived using this redundancy elimination technique offers more fundamental insight. Now the exact locations of solution singularities can be identified without numerical techniques or expensive simulations as was previously done. In addition to more insight, speed improvements on the order of 60x are also realized which allow more particles, and therefore more accuracy per simulation.

1.2 Modeling

1.2.1 History of Electromagnetism

Mankind has always been aware of electromagnetism with the dramatic effects of lightning and only during the past millennium, have we finally begun to gain a complete understanding of the physics behind this fascinating phenomenon. An unconfirmed Greek legend dates the discovery of magnetism back to 900 B.C., when a Greek shepherd named Magnus experienced a force pulling on the iron nails in his sandals while crossing a bed

of black rock [18]. The region was named Magnesia and the black rock became known as magnetite. Electrostatic potential effects were first documented by a Greek philosopher by the name of Thales who noticed that amber could pick up a feather after being rubbed by cat fur. The first widespread use of electromagnetism came around 1100 AD with the development of the compass which relied on using a magnetized needle and the earth's magnetic field for direction.

1.2.2 Coulomb's Law

During the past millennium, hundreds of scientists have contributed to the development of electromagnetic understanding with a plethora of experimentation. One such experimenter, Girolamo Cardano, published De Substilitate (1550) [18] which distinguished for the first time the differences between magnetic and electric forces. One of the first quantitative expressions was later developed by a French physicist, Charles Augustin de Coulomb. Coulomb developed an expression from experimental data that allowed him to predict the force, F, between two charged electrostatic objects [18].

$$F = k_e \frac{q_1 q_2}{r^2} \text{ Coulomb's law} \tag{1.1}$$

The magnitude of the force, F, was directly proportional to the magnitude of a charge, q_1, in the presence of a second charge, q_2, as seen in Fig 1.1. As the two charges q_1 and q_2 were brought closer together, Coulomb noticed that the force was inversely proportional to the distance. The proportionality constant k_e also known as Coulomb's constant was derived from the experimental data but was later shown to have more significance than a simple constant.

1.2.3 Gauss's Law

Despite the usefulness of Coulomb's law, it was not a general expression yet, but valid only for stationary point charges. It was not until 1835 that Carl Friedrich Gauss derived a gerneral expression called Gauss's law [18]. The new expression relied on the concept of an electric field which was first introduced by Michael Faraday. The electric field is defined as the space surrounding an electric charge or moving magnetic field. An electric field represents the force that a hypothetical point charge would experience in a given location. With the new concept of an electric field, Gauss's law stated that the electric flux through any closed surface was proportional to the amount of electric charge enclosed by that surface.

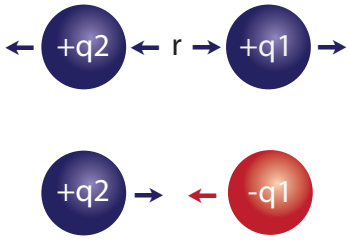

Figure 1.1. Coulomb's law states that the force experienced between two particles q_1 and q_2 is inversely proportional to the distance between them, r^2, and directly proportional to the magnitude of the charges. If charges have equal signs, the force will be a repelling force, and if charges are opposite, the forces will be attractive. Coulomb's law was derived from experimental data compiled by Coulomb. *Figure copyright is owned by author*

The electric flux can be defined as the number of electric field lines moving through an arbitrary surface. Two stronger charges will experience a higher electric flux between the charges since there will be a higher number of equipotential lines. This law can be explained by the expression

$$\text{Integral form: } \oint_S \mathbf{E}\cdot\partial\mathbf{A} = \frac{Q}{\epsilon_0} \tag{1.2}$$

$$\text{Differential form: } \vec{\nabla}\cdot\mathbf{E} = \frac{\rho}{\epsilon_0} \tag{1.3}$$

where ϵ_0 is the material specific permittivity, Q is the total charge, and ρ is the charge density. The differential element, $\partial\mathbf{A}$, is a vector with the magnitude of an infinitesimal surface area element with direction perpendicular to the surface, as seen in Fig 1.2. The divergence[1], $\vec{\nabla}\cdot$, is used to represent the same expression in differential form allowing the benefits of vector calculus to be used.

1.2.4 Gauss's Law of Magnetism

Similar to Gauss's law, an analogous expression was developed for the magnetic field which stated that the total magnetic flux through a closed three-dimensional surface was zero

[1]The divergence is a vector calculus operator which determines the magnitude of a vector field's source or sink. For example, the divergence of the magnetic field, $\tilde{\mathbf{B}}$, can be defined by: $\vec{\nabla}\cdot\vec{B} = \vec{X}\frac{\partial B_x}{\partial x} + \vec{Y}\frac{\partial B_y}{\partial y} + \vec{Z}\frac{\partial B_z}{\partial z}$ for Cartesian coordinates

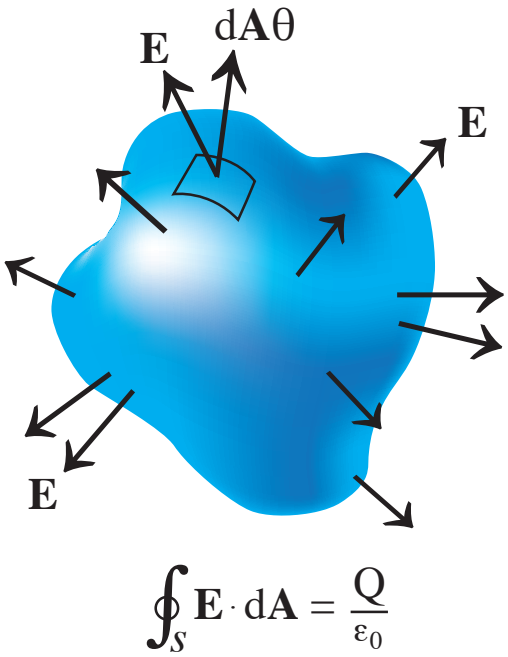

Figure 1.2. Gauss's law states that the electric flux through any closed surface is proportional to the enclosed electric charge. If no charges are contained within the closed surface, no electric flux can move through the surface. If no electric flux moves through the surface, the internal electric field is zero everywhere. *Figure copyright is owned by author*

$$\text{Integral form: } \oint_S \mathbf{B}\cdot\partial\mathbf{A} = 0 \tag{1.4}$$

$$\text{Differential form: } \vec{\nabla}\cdot\mathbf{B} = 0, \tag{1.5}$$

where **B** is the magnetic field and S represents any closed surface enclosing the magnetic field. Unlike Gauss's law for the electric field, magnetic charges are assumed to come in pairs, also known as dipoles. The theoretical existence of a single magnetic point charge, referred to as a magnetic monopole, has been debated with no proof on either side of the argument until recently. New research published in *Nature* September 2009 on spin ice reports quasiparticles resembling the behavior of magnetic monopoles [4]. Gauss's law in the standard set of Maxwell equations (1.8-1.11) excludes the possibility of such a particle existing. Despite this recent discovery, the assumption of no magnetic monopoles has held

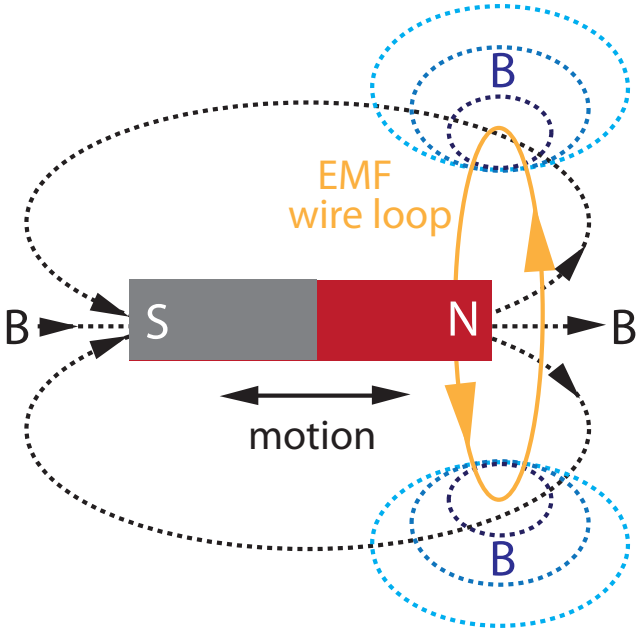

Figure 1.3. Faraday's law of induction states that the electromagnetic force (EMF) around a closed loop is directly proportional to the rate of change of the magnetic flux. The figure demonstrates a simple example where a bar magnet moving through a closed loop induces a current and causes a local magnetic field on the wire loop. This principle is responsible for electricity generated from generators which are capable of harvesting mechanical motion to produce electricity. *Figure copyright is owned by author.*

up experimentally for nearly a century and is considered a valid assumption for the vast majority of cases and therefore is excluded in this work.

1.2.5 Faraday's Law

An additional breakthrough for electrodynamic understanding came from Faraday's law of induction which was also developed from experimental data. Faraday's law of induction states [17] that the induced electromotive force or EMF in any closed circuit is equal to the time rate of change of the magnetic flux ($\frac{\partial \mathbf{B}}{\partial t}$) through the circuit, as shown in Fig 1.3.

$$\text{Faraday's law: } \vec{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{1.6}$$

A similar but opposite law was developed called Ampere's law which related a magnetic field to its current source. In Faraday's law, a moving magnetic flux created a current,

where in Ampere's law, a moving current created a magnetic field.

1.2.6 Ampere's Law

Ampere's law relates the magnetic field outside a closed loop to the current passing through it. The law was derived in 1826 by Andr-Marie Ampre [18]; however, the original formulation did not account for time varying charges. Maxwell added the missing displacement current correction to Ampere's law in 1861 to get

$$\text{Ampere's law: } \vec{\nabla}\times\mathbf{B} = \mu\mathbf{J} + \mu\epsilon\frac{\partial\mathbf{E}}{\partial t} \tag{1.7}$$

where ϵ and μ are material-specific permittivity and permeability, respectively, and $\vec{\nabla}\times$ is the curl[2]. In equation (1.8-1.11) $\mathbf{J}$, ρ and t represent total current density, charge density, and time, respectively.

1.2.7 Maxwell's Equation

In 1864, the greatest single step in the world's knowledge of electricity was made by James Clerk Maxwell [12]. Maxwell, who had compiled the current knowledge of electromagnetic phenomena into twenty differential equations, hypothesized that electric fields traveling at the speed of light was not coincidence. Maxwell's hypothesis that light and electromagnetic fields were related led to his mathematical derivation that light was an electromagnetic wave with oscillating magnetic and electric fields. Maxwell was also able to calculate the speed of light using constants from electrostatic experiments to be 3.107x10$^8 m/s$ (actual speed is: 2.997x10^8). Despite the majority of all current electromagnetic understanding stemming from Maxwell's equations, the general equations that are used today were simplified by Oliver Heaviside [18] to four coupled differential equations:

---

[2]The curl is a vector calculus operator which defines the rotation of a vector field. For example, the curl of the electric field, $\tilde{\mathbf{E}}$, in Cartesian coordinates can be defined by: $\vec{\nabla}\times\mathbf{E} = \vec{X}\left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right) + \vec{Y}\left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}\right) + \vec{Z}\left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right)$

$$\text{Gauss's law: } \vec{\nabla}\cdot\mathbf{E} = \frac{\rho}{\epsilon} \tag{1.8}$$

$$\text{Faraday's law: } \vec{\nabla}\times\mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{1.9}$$

$$\text{Ampere's law: } \vec{\nabla}\times\mathbf{B} = \mu\mathbf{J} + \mu\epsilon\frac{\partial \mathbf{E}}{\partial t} \tag{1.10}$$

$$\text{Gauss's law of magnetism: } \vec{\nabla}\cdot\mathbf{B} = 0 \tag{1.11}$$

Even though Maxwell's equations can be solved directly for simple situations, it is often convenient to introduce the idea of potentials in order to obtain a smaller number of second-order equations. From Maxwell's equations, the electromagnetic expression for light can be derived[7] to be:

$$\left(\vec{\nabla}^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\tilde{\mathbf{E}} = \left(\vec{\nabla}^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\tilde{\mathbf{B}} = 0 \tag{1.12}$$

where c is the speed of light. This expression for light allows us to understand and predict the dramatic effects of light interacting on the nano-scale.

1.2.8 Nanoparticle Arrays

Recent attention in nanoscience has focused on the potential uses found in periodic arrays of metal nanoparticles. Silver (Ag) and Gold (Au) arrays have been shown to have dramatic sensitivities to environmental refractive index changes making them ideal candidates for biosensors. This work focuses specifically on the experimental and theoretical results of spherical Au arrays, as shown in Fig 1.4. Using similar nanoparticle arrays, single cell analyte detection is possible due to rapid changes in the elctromagnetic behavior of such arrays.

1.2.9 Nanoparticle Array Modeling

In §2.1-2.4, theoretical methods are discussed for predicting the behavior of light with the nanoparticle (NP) arrays. Numerical methods are available for solving the coupled partial differential methods which make up Maxwell's equation. However, high computational costs usually limit a direct approach due to the number of particles and the standard numerical methods available which require spacial discretization. Other numerical methods which are faster and more commonly used to model the behavior of light interacting with NP arrays include Finite Difference Time Domain (FDTD) [19], T-matrix [14], and

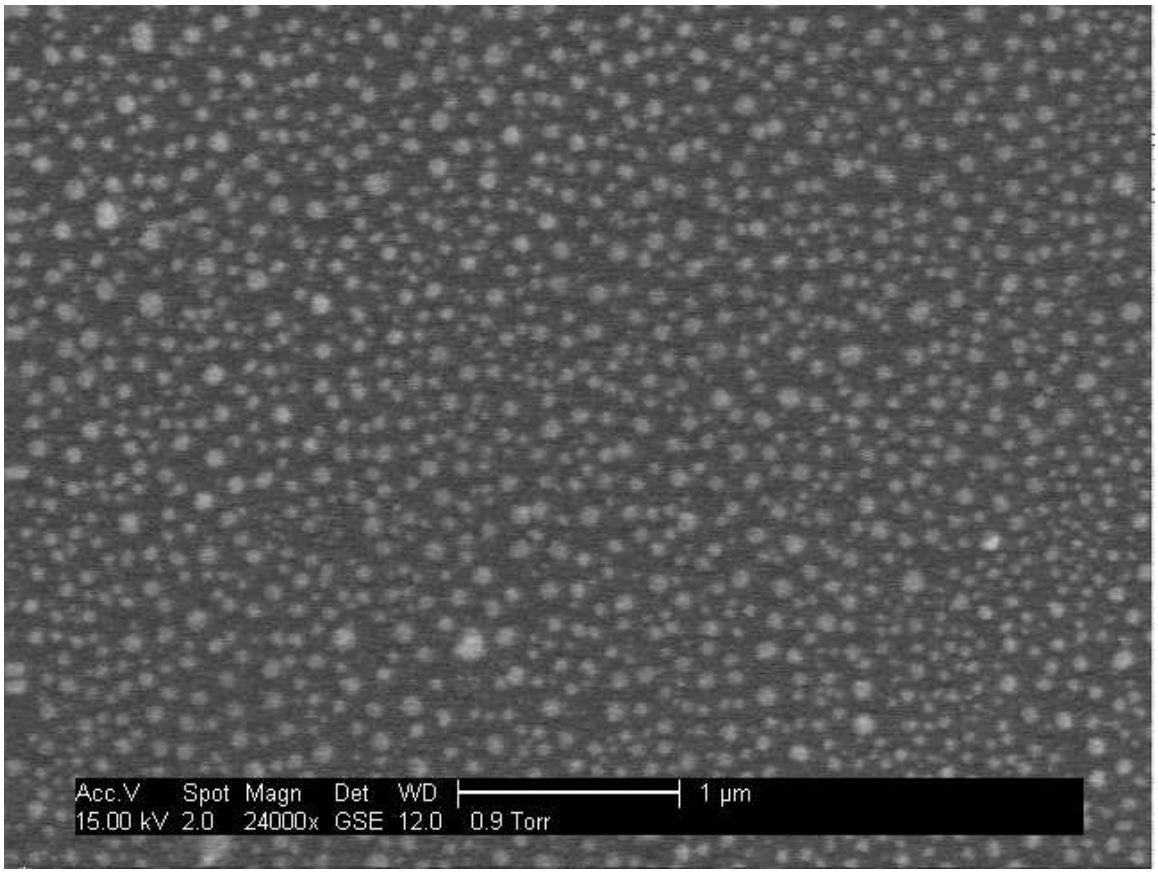

Figure 1.4. SEM image of a randomly oriented Au spherical nano particle array. The array is periodic, and therefore continues randomly across the entire chip. The chemical methods used to produce this sample are outlined in detail in Chapter 3. *Reprinted with permission from Langmuir [23]*

Coupled Dipole Approximation (CDA) [21]. FDTD and T-matrix are considered exact multipole methods because they account for all significant poles. The CDA, which is only a dipole approximation, can be valid for smaller NP sizes $\leq$100 nm. Despite the relative computational speed improvement using the CDA approach outlined by Schatz [22], optimization of design parameters is still limited. A new analytical solution derived in this work can now provide precise locations for all singularities for a square grid array for particle sizes $\leq$100 nm for nonmagnetic particles. Singularity design criteria are of particular interest because they are usually related with enhanced sensitivity characteristics. The new solution, derived in Section §2.2, provides more fundamental insight and the computational efficiency necessary to allow true multivariable optimization of all design parameters: NP spacing, NP size, NP material, medium material, and wavelength used.

1.3 Fabrication

1.3.1 Nanoparticle Array Fabrication

Chapter 3 introduces fabrication of Au NP arrays on Silicon (Si) substrate. Various methods of Au NP array fabrication have been developed including ion implantation, surface etching [16], laser ablation [9], inductively coupled thermal plasmas [11], thermodynamic self organization [8], and thermal evaporation of thin Au film followed by annealing [6]. In addition to NP array fabrication, Au thin films and island films (thin films with voids) have been created using electrodeposition, electroless deposition (EL) [15], sputtering [10], and vacuum evaporation [5]. Current methods for fabrication tend to be limited by geometry, material, and cost. In collaboration with Ahn et al. [2] a *new* EL method has been developed, as outlined in Chapter 3, which allows for tunable NP size distributions as well as the ability to coat internal 3D surfaces at ambient conditions. This new EL method is also preferred since it does not require expensive equipment and has the ability for rapid coating and fabrication making it a potential economical candidate for large scale NP array fabrication.

1.4 Analysis

1.4.1 Scanning Electron Microscope

A Back Scattered Electron (BSE) SEM is popular for nano-scale analysis. It works by using a high-powered electron beam that emits electrons down onto the sample. The sample is typically placed in a high-vacuum environment in order to limit additional noise that an atmosphere would induce. The electron beam is moved with angstrom ($1x10^{-10}$ meters) precision using magnetic fields to focus and scan the beam in a raster (*side-to-side*) pattern. As shown in Fig 1.5, while the beam is scanning, electrons from the beam are scattered and collected by a BSE detector.

Within the detector, the electrons enter a photomultiplier tube to amplify the signal which is then converted into an image. The SEM is susceptible to image noise and therefore, additional steps are needed which include averaging multiple scans and applying digital filters [3]. BSE SEM offers a powerful tool capable of obtaining topographical sample dimensions, material composition, and electrical conductivity.

1.4.2 Atomic Force Microscopy

In addition, the AFM [1] has been developed to get topography information on samples using a mechanical cantilever arm that is moved across the surface in a raster pattern; see

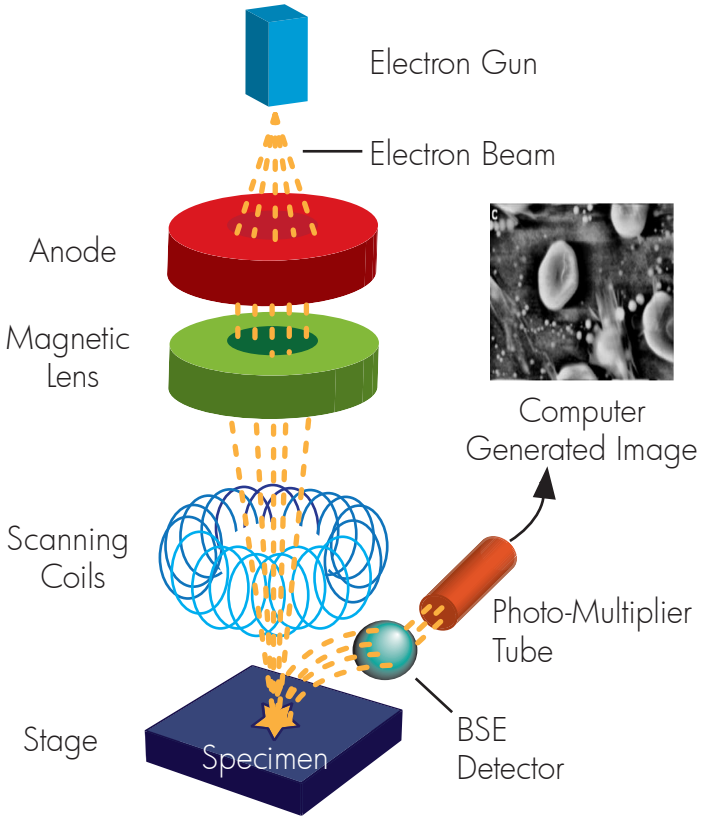

Figure 1.5. Scanning Electron Microscope (SEM) diagram showing the basic elements of a SEM. The electron gun bombards the sample with an electron beam and is focused and moved with magnetic lenses and fields. The sample is typically placed within a high vacuum environment in order to limit the additional noise that an atmosphere induces. The primary electrons from the beam are scattered according to the topography and material of the sample. An electron collector collects the scattered electrons and amplifies the signal with a photomultiplier tube. The resulting signal is converted into a raster image as the electron beam scans the sample side-to-side. [23] *Figure copyright is owned by author*

Fig 1.6. As the microscopic lever moves up and down over the particles or changes in the sample topography, an electrical signal is detected via a laser and converted into an image. The AFM is considerably slower than the SEM due to the mechanical movement of the arm. The AFM also lacks resolution relative to the SEM. However, the AFM does offer topographic data that can complement the SEM analysis. A detailed comparison between the data sets offered with SEM and AFM is presented later in this work in Chapter 4.

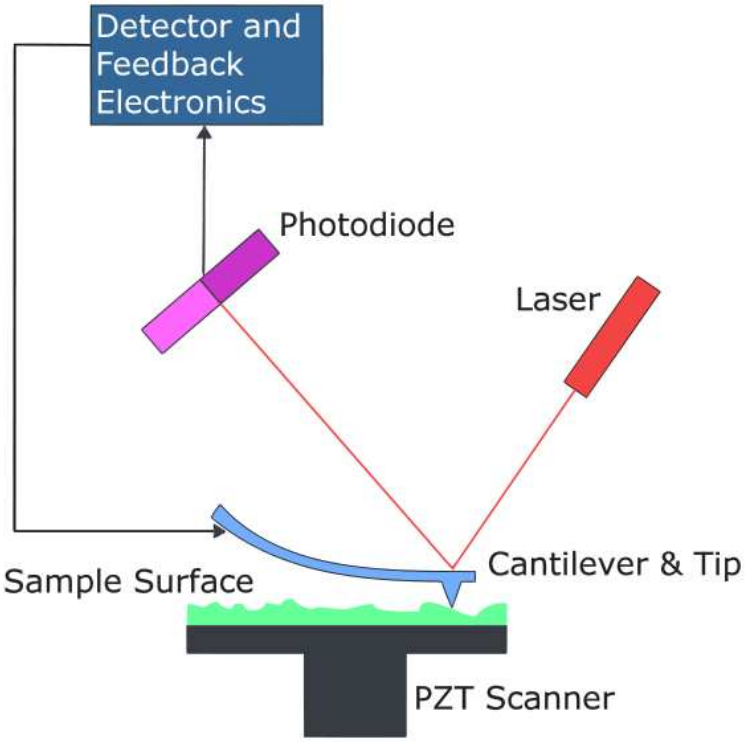

Figure 1.6. Atomic Force Microscope (AFM) diagram showing basic elements of an AFM. A cantilever tip moves along the top of the sample in order to detect sample topography. As the cantilever moves up and down, a laser reflects off the cantilever arm to detect movement via a photo diode. A raster image is made from the electrical signal detected by the photo diode as the arm is scanned side-to-side. *Reprinted with permission from Wikipedia [3]*

1.4.3 Nanoparticle Sample Analysis

Chapter 4 reviews common methods for NP data acquisition and related image processing algorithms used for analysis. Popular tools for NP data acquisition include AFM and SEM images which have been previously mentioned. However, AFM tip convolution produced by the AFM probe tip can cause exaggerated particle diameters if improper algorithms are used. A detailed comparison between the data produced by AFM and SEM is given in Chapter 4 to provide a better understanding between these two methods. Once the data have been collected, it can be analyzed using various image processing algorithms. Commonly used algorithms include threshold particle segmentation, which uses a particular gray scale shade to designate particles from silicon. The threshold method is simple but can also create exaggerated particle sizes due to clumping multiple particles into one. Superior methods include using the watershed transform [13] which is capable of segmenting adjacent

particles. In addition to particle segmentation, an additional algorithm has been developed which allows particle shape analysis (see 4.1.1).

1.5 Conclusion

Making use of thousands of years of electromagnetic understanding leading to Maxwell's equation, all macroscopic electrodynamic phenomena can be explored. Individual equations and laws previously derived have been scrutinized and re-derived to produce a sound set of equations that have held up to experimental data. Exciting enhancements at the nano scale can be modeled using computational approximation and in some cases analytical solutions derived from Maxwell's equations for a greater understanding. In addition to modeling, Au NP arrays can be fabricated and analyzed to provide greater understanding, theoretical validation, and feedback on model assumptions. The Greek prefix *nano* is becoming more well known as researchers and scientists break new ground in the nanoscience area using powerful tools of function call reduction discussed in this work.

CHAPTER 2

THEORY AND MODELING

2.1 PNAS Submission

This chapter was submitted to the Proceedings of the National Academy of Science.

Exact solution to the coupled dipole approximation of Maxwell's equations for electromagnetic interactions in periodic nanostructures

Ben J. Taylor*, Braden M. Harbin†, and D. Keith Roper*†

*University of Utah Department of Chemical Engineering, and †University of Arkansas Ralph E. Martin Department of Chemical Engineering

Submitted to Proceedings of the National Academy of Sciences of the United States of America

An exact solution to the coupled dipole approximation of Maxwell's equations for electromagnetic interactions in periodic subwavelength nanostructures has been derived. Contributions to the retarded dipole sum from near-field, intermediate, and far-field radiative interactions are distinguished. The solution precisely identifies the source, location, and magnitude of periodic divergences in real and imaginary components of the retarded dipole sum which arise due to the wavenumber and grid size, independent of particle size and dielectric constants of particle and environment. Far-field radiative contributions are significant, outweighing near-field and intermediate effects. Real and imaginary parts computed for retarded dipole sum in finite arrays converge toward the exact solution as array size increases. The exact solution permits validation of approximations to retarded dipole sum using finite difference time domain and discrete dipole approaches. The new solution allows rapid optimization and design of novel optoelectronic, spectroscopic, biomedical and sensing systems which use nanoparticle arrays as well as integrated circuit waveguides, optical filters, photovoltaics and antennae.

Section 1: Introduction

Periodic subwavelength nanostructures such as one- and two-dimensional arrays of holes or particles exhibit extraordinary spectral resonances in theory [1, 2, 3, 4, 5, 6, 7, 8] and experiment [9, 10, 11, 12, 13, 14, 15]. Periodic nanostructures allow superluminal propagation [16] and could improve surface enhanced spectroscopies [17, 18, 19], sensors [20, 21], photodynamic therapies [22], subdiffractive mapping of biomolecules [23], integrated circuit waveguides [24], optical filters, photovoltaics and antennae. Lattice dimension [7] and material size, shape, and dielectric functions can be selected to phase-match incident and diffractive coupled fields and induce large polarizabilities that enhanced local surface fields and produce narrow, high-amplitude resonances at tunable frequencies. Control of strong local EM fields has recently permitted nanoparticle (NP) tuning of magnetic resonance frequency and lineshape of split ring resonator metamaterials [25] and modulation of polarization of light scattered from NP [26].

Rapid development of spectroscopically-active periodic subwavelength nanostructures has focused recent attention on solving Maxwell's equations for electromagnetic (EM) interactions in nanoscale lattices. The complete problem has been analyzed computationally using discrete dipole approximation (DDA), finite difference time domain (FDTD), and T-matrix. Some limiting cases of the problem have been solved analytically to obtain exact expressions. However, existing exact expressions for two-dimensional (2D) arrays neglect contributions to local electromagnetic field from far-field radiation. Solutions to classical EM theory are necessary, to describe optical properties of multi-NP systems such as surface enhanced Raman spectroscopy (SERS) [27] especially in combination with quantum mechanical methods for electric fields very close to NP surfaces [28]. Exact solutions improve matching of atomistic and continuum-scale methods to allow multiscale characterization of such phenomena.

The DDA [29, 30] allows scattering and absorption of an incident EM field to be computed for dielectric structures of arbitrary geometry and composition from induced dipole moments that arise from interactions in a set of point dipoles representing a scatterer. DDA results for nanoscale lattices have been validated using exact solutions to Maxwell's equations for two contiguous spheres [31]. DDA results were comparable with T-matrix (i.e. coupled multipole) simulation [24] of scattering from arrays based on Mie theory using a truncated multipolar expansion around each particle and the translation addition theorem for vector harmonics [3]. DDA accurately predicted spectral wavelength shifts of Ag cylinders [11].

Compared with FDTD, DDA computations are generally applicable, easier to use, faster, more accurate and more sensitive to local dielectric constant. But computational cost becomes prohibitive when solving the number of equations required for large arrays, and when varying multiple parameters (e.g. particle size, spacing, and dielectric constant) [31]. DDA computation using the package DDSCAT [17, 18, 19] inaccurately predicted values of polarization in large particle arrays for interparticle spacing equal to the wavelength of light λ [8]. While some numerical methods could take advantage of symmetry in the scatterer, symmetry has not been used to simplify DDA computation.

The coupled dipole approximation (CDA) is formally similar to DDA, except that each scatterer is represented by a single dipole whose polarizability is equal to that of the particle. A lattice dispersion formula may be used to make the dipole array response identical to that of the bulk continuum. T-matrix provides exact solutions to the CDA that allow attribution and quantification of resonant-energy transmission and absorption features. These features are determined by material dielectric properties, nanostructure dimensions and periodicity on the basis of limiting condition sets. Computational results that used

Reserved for Publication Footnotes

CDA for infinite arrays of NP were shown to produce spectral features consistent with those from T-matrix for interparticle spacings, D, greater than $2R$ in NP of radius, R, equal to 30nm and 5 nm, respectively [3].The adequacy of CDA may decrease in describing NP $>$ 100 nm in which multipolar dynamic depolarization and radiative damping decrease scattering.

Exact expressions for induced dipole fields in periodic arrays of infinite dimensions have been derived for one-dimensional (1D) dipole chains [2] and for 1D chains of cylinders [8] in the long-wavelength limit ($r_{ij} \ll \lambda$) and for 2D dipoles *neglecting* far-field radiation, where r_{ij} is the spacing between between two particles i and j. Merchiers [32] derived expressions for far-field cross sections of 2D arrays as functions of induced electric and magnetic dipole moments. Comparing analytic dipole and computational multipole approaches shows multipole effects produce large deviations in spectral peakshifts only for $|r_{ij}|/R < 2.5$ [33]. This validates development of exact expressions that neglect multipole effects in this range.

Experimental studies of periodic subwavelength nanostructures has typically been limited to disk shaped NP arrays made by electron-beam lithography [34]. This is due to difficulties in fabricating stable arrays of spheres. These studies show increasing interparticle spacing (i.e. grating constant, $|r_{ij}|$) increases transverse-mode resonance peak and scattering intensity while decreasing Lorentzian peakwidth. This increases Q-factor due to ohmic loss from electron-phonon scattering that occurs instead of radiative loss.

The purpose of this work has been to derive an exact solution to the CDA of Maxwell's equations for electromagnetic interactions in periodic dipole arrays that include contributions due to far-field radiation. This exact solution provides a benchmark for resource-intensive computational analysis. It is applicable to analyze superluminal propagation [2]. An exact solution makes it possible to extend results recently computed from Maxwell's equations via generalized Mie theory (GMT) [26, 27], DDA [28], and FDTD [22] in order to describe modulation of polarization, enhanced EM fields (SERS), and photodynamic therapy, respectively, and concepts such as NP sensing networks [20, 21], light-induced self assembly [35], and subdiffractive mapping of biomolecules [23] examined in small, random NP clusters to regular, two-dimensional NP systems. It allows interpretation of phenomena observed in physical subwavelength nanostructures of increasing complexity. It can be used to guide parameter selection in design and manufacture of integrated devices. Examples summarized in this and previous paragraphs represent some of many possible specific applications for the exact solution.

Section 2: Model

The phenomena of classical electricity and magnetism interacting with physical objects can be completely described on a macroscopic scale using Maxwell's equations and the Lorentz Force law, $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, where $\mathbf{F}$ is the force experienced by a particle, q is the charge on the particle, $\mathbf{E}$ is the electric field, $\mathbf{v}$ is the particle velocity, and $\mathbf{B}$ is the magnetic field. In a homogeneous, isotropic, linear medium in which magnetic monopoles are absent the Maxwell equations become

$$\vec{\nabla}\cdot\mathbf{E} = \frac{\rho}{\epsilon} \quad [1]$$

$$\vec{\nabla}\times\mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad [2]$$

$$\vec{\nabla}\times\mathbf{B} = \mu\mathbf{J} + \mu\epsilon\frac{\partial \mathbf{E}}{\partial t} \quad [3]$$

$$\vec{\nabla}\cdot\mathbf{B} = 0 \quad [4]$$

where ϵ and μ are material-specific permittivity and permeability, respectively, and $\vec{\nabla}\times$ is the curl. In (1-4) $\mathbf{J}$, ρ and t represent total current density, charge density, and time respectively. Even though Maxwell's equations can be solved directly for simple situations it is often convenient to introduce the idea of potentials in order to obtain a smaller number of second-order equations. For the case of nonmagnetic material we can neglect[1] the magnetic field Maxwell equations. The electric field may be defined in terms of a vector potential, $\mathbf{A}$, when there are no sources ($\rho = 0$ and $\mathbf{J} = 0$) using

$$\mathbf{E} = \frac{ic}{k}\vec{\nabla} \times (\vec{\nabla} \times \mathbf{A}) \quad [5]$$

where $k = \omega\sqrt{\mu\epsilon}/c = 2\pi/\lambda$ is the wavenumber, ω represents angular frequency, λ is wavelength, c is the speed of light, and $i = \sqrt{-1}$. In Jackson [37] the sinusoidal time varying retarded vector potential is defined as

$$\mathbf{A}(\mathbf{x}) = \frac{\mu}{4\pi}\frac{e^{ikr}}{r}\sum_{m=0}^{\infty}\frac{(-ik)^m}{m!}\int \mathbf{J}(\mathbf{x}')(\mathbf{n}\cdot\mathbf{x}')^m d^3x'$$

[6]

where $\mathbf{x}$ is the point where the field is measured, $\mathbf{x}'$ corresponds to the source of the $\mathbf{E}$ field, and $\mathbf{n}$ is the unit vector between the points $\mathbf{x}$ and $\mathbf{x}'$.

The vector potential for a dipole can be derived by keeping only the 0th term (m=0) in (6) to produce

$$\mathbf{A}(\mathbf{r}) = -\frac{\mu}{4\pi}\frac{e^{ikr}}{r}i\omega\mathbf{p} \quad [7]$$

[1] The effective μ may deviate from unity [36] when the second Mie coefficient, b_1, is no longer negligible for nonmagnetic materials (e.g. Au, Ag, ...).

where $\mathbf{r} = |\mathbf{x} - \mathbf{x}'|$ is the displacement vector and $\mathbf{p}$ is the electric dipole moment defined by $\mathbf{p} = \int \mathbf{x}'\rho(\mathbf{x}')d^3x'$, and $r = |\mathbf{r}|$. The dipole vector potential can be substituted into the electric field defined in terms of a vector potential to produce [*see Supporting Information: Appendix A*]

$$\mathbf{E}(\mathbf{r},\mathbf{p}) = \frac{e^{ikr}}{4\pi\epsilon}\left\{\left(\frac{1}{r^3} - \frac{ik}{r^2}\right)[3(\hat{\mathbf{r}}\cdot\mathbf{p})\hat{\mathbf{r}} - \mathbf{p}] + \frac{k^2}{r}[(\hat{\mathbf{r}}\times\mathbf{p})\times\hat{\mathbf{r}}]\right\} \quad [8]$$

which is the general vector form of the electric field of a retarded oscillating dipole where $\hat{\mathbf{r}}$ is the unit vector in the $\mathbf{r}$ direction, defined by $\hat{\mathbf{r}} = \mathbf{r}/|\mathbf{r}|$.

We now proceed to use the definition of the dipole electric field (8) to model the behavior of light interacting with a nanoparticle (NP) array. The local electric field, $\mathbf{E_{loc}}$, experienced by an individual NP upon which light is incident is made up of a contribution due to the oscillating electric field, $\mathbf{E}$, from neighboring NPs given by (8) and a contribution due to incident light, $\mathbf{E_{inc}}$. The electric field due to the incident light can be derived from Maxwell's equations when there are no free sources and ϵ and μ are assumed constant to produce a wave equation,

$$\vec{\nabla}^2\mathbf{E} - \mu\epsilon\frac{\partial^2\mathbf{E}}{\partial(ct)^2} = 0 \quad [9]$$

whose solution is

$$\mathbf{E_{inc}} = \mathbf{E_o}e^{i\mathbf{k}\cdot\mathbf{r}'-i\omega t} \quad [10]$$

where $\mathbf{r}'$ is a position in the wavefield and $\mathbf{k}$ is the wavevector. Using the principle of superposition the local electric field experienced by any individual NP can be expressed as

$$\mathbf{E_{loc}} = \mathbf{E_{inc}} + \sum_{j\neq i}\mathbf{E}(\mathbf{r}_{ij},\mathbf{p}_j) \quad [11]$$

where the subscripts i and j denote a vector between the origin, i, and the neighboring particle, j.

The expression for the local electric field can be used to relate polarization, $\mathbf{p_i}$, to the local field using a scalar polarizability, α_i, viz.

$$\mathbf{p_i} = \alpha_i\mathbf{E_{loc}}. \quad [12]$$

For the case when the incident light is perpendicular to the array in the $\hat{\mathbf{x}}$-$\hat{\mathbf{y}}$ plane and polarized in the $\hat{\mathbf{x}}$ direction and the array is symmetric about both $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ (8) can be used to produce [*see Supporting Information: Appendix A*]

$$\sum_{j\neq i}\mathbf{E}(\mathbf{r}_{ij},\mathbf{p}_j) = \sum_{j\neq i}\frac{e^{ikr_{ij}}}{4\pi\epsilon}\left\{\left(\frac{1}{r_{ij}^3} - \frac{ik}{r_{ij}^2}\right)(3\cos^2(\theta_{ij}) - 1) + \frac{k^2\sin^2(\theta_{ij})}{r_{ij}}\right\}\mathbf{p_j} \quad [13]$$

where θ_{ij} is the angle between the position vector, $\hat{\mathbf{r}}_{\mathbf{ij}}$, and the polarization, $\mathbf{p_j}$, and $r_{ij} = |\mathbf{r_{ij}}|$.
Now after combining (13), (11) into (12) we obtain

$$\mathbf{p_i} = \alpha_i\mathbf{E_{inc}} + \frac{\alpha_i}{4\pi\epsilon}\sum_{j\neq i}e^{ikr_{ij}}\left\{\overbrace{\frac{3\cos^2(\theta_{ij})-1}{r_{ij}^3}}^{static\ (SZ)} - \overbrace{\frac{ik(3\cos^2(\theta_{ij})-1)}{r_{ij}^2}}^{intermediate\ (IZ)} + \overbrace{\frac{k^2\sin^2(\theta_{ij})}{r_{ij}}}^{radiation\ (RZ)}\right\}\mathbf{p_j}. \quad [14]$$

In (14) contributions to the oscillating dipole electric field arise from near-field static zone (SZ;$1/r^3$), induction zone (IZ;$1/r^2$), and far-field radiation zone (RZ;$1/r$). In the case of an infinite array, $\mathbf{p_i} = \mathbf{p_j} = \mathbf{p}$ and (14) can be solved for $\mathbf{p}$ yielding

$$\mathbf{p} = \alpha\mathbf{E_{loc}} = \frac{\mathbf{E_{inc}}}{\frac{1}{\alpha} - S(k)} = \frac{\mathbf{E_o}e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t}}{\frac{1}{\alpha} - S(k)} \quad [15]$$

where $S(k)$ is the retarded dipole sum given by:

$$S = \frac{1}{4\pi\epsilon}\sum_{\substack{j=1\\j\neq i}}^{\infty}f_{i,j}(k) = \frac{1}{4\pi\epsilon}\sum_{\substack{j=1\\j\neq i}}^{\infty}\left\{\overbrace{\frac{3\cos^2(\theta_{ij})-1}{r_{ij}^3}}^{SZ} - \overbrace{\frac{ik(3\cos^2(\theta_{ij})-1)}{r_{ij}^2}}^{IZ} + \overbrace{\frac{k^2\sin^2(\theta_{ij})}{r_{ij}}}^{RZ}\right\}e^{ikr_{ij}}. \quad [16]$$

The retarded dipole sum, $S(k)$, corresponds to the sum of the electric dipole fields at any array node divided by the polarization, $\mathbf{p}$. The local electric field, $\mathbf{E_{loc}}$, in the infinite array is equivalent at every array node. Its magnitude and direction varies periodically from node to node throughout the infinite array.

The function, $f_{i,j}(k)$, in (16) can be transformed to a function $f_{i,j}(k,D)$ in rectilinear coordinates using $r_{ij} = D\sqrt{x^2+y^2}$ and angle $\theta_{ij} = \tan^{-1}(y/x)$. In this transformation, variables x and y are integer multiples of intraparticle spacing, D, for a square array along the corresponding coordinate axes $\hat{x}$ and $\hat{y}$ (See Fig 1).

The transformation for θ_{ij} and r_{ij} reflects a case when the polarization is parallel to the rows of particles in the $\hat{x}$ direction and $\mathbf{k}$ is perpendicular to the array. After substitution into (16) we obtain

$$f_{x,y}(k,D) = \left\{ \frac{\overbrace{(2x^2-y^2)}^{SZ}}{D^3\sqrt{x^2+y^2}^5} + \frac{\overbrace{ik(y^2-2x^2)}^{IZ}}{D^2(x^2+y^2)^2} + \frac{\overbrace{k^2y^2}^{RZ}}{D\sqrt{x^2+y^2}^3} \right\} e^{ikD\sqrt{x^2+y^2}}. \quad [17]$$

For selected values of variable k and parameter D, the function $f_{x,y}$ exhibits symmetry around both $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ axes. Figure 1 shows that by choosing to center the reference frame within the infinite lattice we may take advantage of mirror symmetry about the $\hat{x}$ and $\hat{y}$ axes. We select this coordinate system to reduce the number of function calls to compute S. The computation of $f_{x,y}$ for any dipole in quadrant II, III, and IV thereby reduces to computation of the corresponding symmetric point of $f_{x,y}$ in quadrant I. We later introduce an algorithm to explicitly identify these symmetric points.

To ensure a single center point exists we increment the grid size, L, by odd numbers. The retarded dipole sum, S, may be approximated using a finite grid size $L = 2N+1$, $N = 1, 2, 3 \ldots \infty$, to obtain $S^{L\times L}$. To identify symmetric $f_{x,y}$ points which are redundant computations in the truncated S^{LxL} summation we visualize individual summation terms for a given k and D as matrix elements. For $N = 1$ and $N = 2$, respectively, we obtain

$$S^{3\times 3} = \frac{1}{4\pi\epsilon} \sum \begin{bmatrix} f_{1,1} & f_{0,1} & f_{1,1} \\ f_{1,0} & 0 & f_{1,0} \\ f_{1,1} & f_{0,1} & f_{1,1} \end{bmatrix} \text{ and} \quad [18]$$

$$S^{5\times 5} = \frac{1}{4\pi\epsilon} \sum \begin{bmatrix} f_{2,2} & f_{1,2} & f_{0,2} & f_{1,2} & f_{2,2} \\ f_{2,1} & f_{1,1} & f_{0,1} & f_{1,1} & f_{2,1} \\ f_{2,0} & f_{1,0} & 0 & f_{1,0} & f_{2,0} \\ f_{2,1} & f_{1,1} & f_{0,1} & f_{1,1} & f_{2,1} \\ f_{2,2} & f_{1,2} & f_{0,2} & f_{1,2} & f_{2,2} \end{bmatrix} \quad [19]$$

where each element in the respective $L \times L$ matrix of (18) and (19) are summed to obtain the truncated estimator, $S^{L\times L}$, of the dipole sum. Computation of $S^{L\times L}$ is simplified by considering the symmetry, which reduces the number of function calls, $(N+1)^2 - 1$, relative to the number of elements in the matrix, $2N+1$, by a fraction of

$$\frac{(N+1)^2 - 1}{(2N+1)^2} \quad [20]$$

to provide a first estimate to the improved speed of the new algorithm. For large grid sizes, i.e. $N \to \infty$, (20) approaches 1/4.

Once redundant matrix elements have been identified they can be factored out and grouped. This can be done by writing the subscript variables x and y of $f_{x,y}$ into two M-dimensional vectors $\mathbf{X}$ and $\mathbf{Y}$, where $M = 8\sum_{i=1}^{N} i$. This process may begin at the top center position of the outer shell in an $L \times L$ matrix and move clockwise around the perimeter, repeating the process for each consecutive nested shell in the matrix to fill the vectors. For example, in (18) the subscript variables of the top center function call $f_{0,1}$ are transformed into initial elements of the respective vectors, $\mathbf{X}_1 = 0$, and $\mathbf{Y}_1 = 1$. Repeating this process for the entire matrix in (18) yields

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix} \quad [21]$$

With the input variables of the 3×3 array expressed in vector form (21) it is clear that only three function calls are required to produce results for all eight elements in the array. For larger non-square or multidimensional arrays we introduce an algorithmic approach to group the redundant function calls. Such an algorithm would be helpful to eliminate redundant function calls for 3D clusters of NPs. An algebraic grouping algorithm is helpful since visual grouping in more complex situations becomes tedious. The truncated estimator, $S^{L\times L}$, can thus be represented in a linear algebraic format after grouping identical function calls in (21), viz.,

$$S^{3\times 3} = \frac{1}{4\pi\epsilon} \left(4f_{1,1} + 2(f_{1,0} + f_{0,1})\right). \quad [22]$$

The resulting expression, (22), factors out all apparent redundant function calls of (18) thereby speeding computation. For larger arrays the redundant function calls can likewise be factored out to yield for N=2 and N=3

$$S^{5\times 5} = S^{3\times 3} + \frac{1}{4\pi\epsilon} \left(4(f_{2,1} + f_{2,2} + f_{1,2}) + 2(f_{2,0} + f_{0,2})\right) \quad [23]$$

$$S^{7\times 7} = S^{5\times 5} + \frac{1}{4\pi\epsilon} \left(4(f_{3,1} + f_{3,2} + f_{3,3} + f_{1,3} + f_{2,3}) + 2(f_{3,0} + f_{0,3})\right) \quad [24]$$

respectively. We group like terms in (22), (23), and (24) to reveal a trend which allows the function calls to be represented generally for any value of L:

$$S^{L\times L} = S^{(L-2)\times(L-2)} + \frac{1}{4\pi\epsilon}\left\{2(f_{N,0}+f_{0,N}) + 4\left(\sum_{j=1}^{N} f_{N,j} + \sum_{j=1}^{N-1} f_{j,N}\right)\right\}. \quad [25]$$

The new expression, (25), may be simplified further by combining the summation terms into a single summation. This is done by adding an additional term, $f_{N,N}$, to the overall expression in (25) resulting in:

$$S^{L\times L} = \overbrace{S^{(L-2)\times(L-2)}}^{interior} + \frac{1}{4\pi\epsilon}\left\{\overbrace{2(f_{N,0}+f_{0,N}) + 4f_{N,N} + 4\sum_{j=1}^{N-1}[f_{N,j}+f_{j,N}]}^{perimeter}\right\}. \quad [26]$$

The leftmost term on the right hand side of (26) represents the interior points on any $L \times L$ grid. The bracket terms in (26) represent the perimeter points that remain to be calculated. Now (26) can be rewritten as a double summation to account for the nesting of consecutive shells in the $S^{(L-2)\times(L-2)}$ term that make up the interior points.

$$S^{L\times L} = \frac{1}{4\pi\epsilon}\sum_{m=1}^{N}\left[\overbrace{2(f_{m,0}+f_{0,m})}^{axial} + \overbrace{4f_{m,m}}^{diagonal}\right] + \overbrace{\frac{1}{\pi\epsilon}\sum_{m=2}^{N}\sum_{n=1}^{m-1}[f_{m,n}+f_{n,m}]}^{off-axial/diagonal(OAD)} \quad [27]$$

where m and n are integers used to increment inner and outer summations. The double summation term in (27) includes all of the off-axial/diagonal (OAD) elements. In (27) the axial elements are represented by $f_{m,0}$ and $f_{0,m}$ and the diagonal elements are represented by $f_{m,m}$. The remaining terms $f_{m,n}$ and $f_{n,m}$ represent the remaining elements that are off-axial and off-diagonal. With the new algorithm, an expression similar to (20) can be derived to estimate the overall improvement in computing speed. The number of function calls, $((N+1)(N+2)-1)/2$, relative to the number of elements in the matrix is represented by the fraction

$$\frac{(N+1)(N+2)-1}{2(2N+1)^2} \quad [28]$$

where for large grid sizes ($N \to \infty$), (28) approaches 1/8 for the new algorithm.

The symbolic functions in (17) are now substituted into the new algorithm in (27), which consists of nested double-sum function calls, to allow algebraic and trigonometric simplification of function terms, using a reciprocal scaled wavelength defined as $A = kD = \frac{2\pi D}{\lambda}$. Its use ultimately permits single variable interpolation. This reduces (16) when $N = \infty$, to the following solution [*see Supporting Information: Appendix B*]:

$$S(A,D) = \frac{1}{4\pi\epsilon D^3}(2S_I(A) + S_{II}(A) + 4S_{III}(A)), \quad [29]$$

$$S_I(A) = \overbrace{P_3(e^{iA})}^{axial\ SZ} - \overbrace{iAP_2(e^{iA})}^{axial\ IZ} - \overbrace{A^2\ln(1-\cos(A)-i\sin(A))}^{axial\ RZ} \quad [30]$$

$$S_{II}(A) = \overbrace{\frac{P_3(e^{i\sqrt{2}A})}{\sqrt{2}}}^{diag\ SZ} - \overbrace{iAP_2(e^{i\sqrt{2}A})}^{diag\ IZ} - \overbrace{\sqrt{2}A^2\ln(1-\cos(\sqrt{2}A)-i\sin(\sqrt{2}A))}^{diag\ RZ} \quad [31]$$

$$S_{III}(A) = \overbrace{Q_3(e^{iA})}^{OAD\ SZ} - \overbrace{iAQ_2(e^{iA})}^{OAD\ IZ} + \overbrace{A^2Q_1(e^{iA})}^{OAD\ RZ} \quad [32]$$

where P_m is a polylogarithm defined by

$$P_m(Z) = \sum_{n=1}^{\infty}\frac{Z^n}{n^m} = \int_0^Z \frac{P_{m-1}(t)}{t}dt \quad [33]$$

and $Q_z(x)$ is a double sum defined as

$$Q_z(x) = \sum_{m=2}^{\infty}\sum_{n=1}^{m-1}\frac{x^{\sqrt{m^2+n^2}}}{(m^2+n^2)^{z/2}} \quad [34]$$

and can be used to understand and simplify (32). The polylogarithms are special functions that can quickly be calculated within machine precision using the integral definition, and for which analytical solutions exist for various cases. The complete solution, $S(A,D)$, includes near- and far-field contributions due to interactions between axial ($S_I(A)$), diagonal ($S_{II}(A)$), and off-axial/off-diagonal ($S_{III}(A)$) elements. The latter semi-analytical component requires further numerical reduction [*see Supporting Information: Appendix C*].

Section 3: Discussion

Divergences: existence and location. The new solution in (29) - (32) provides fundamental insight into the extinction spectra and increases computational speed and accuracy. The product $S(A,D)D^3$ is a function of only the reciprocal scaled wavelength, $A = kD$. Singularities and divergences in exist when

$$A = kD = \frac{2\pi n}{\sqrt{x_f^2 + y_f^2}} \quad [35]$$

where the subscript f represents the first element along axial and diagonal rays of particles emanating from i at a constant angle, θ_{if}, and n is an integer multiple. For example, in the case of (30) when $x_f = 0$ and $y_f = 1$, (35) produces $A = 2\pi n$ which when substituted into (30) produces divergences when the natural log goes to zero for all n. Likewise, in the case of (31) when $x_f = 1$ and $y_f = 1$, (35) produces $A = 2\pi n/\sqrt{2}$ which when substituted into (31) produces divergences for all n. The expression in (35) is equivalent to saying divergences exist when the wavelength is equal to the distance between the center particle i and the first element f,

$$\lambda = \frac{r_{if}}{n}. \quad [36]$$

Real and imaginary solution components: axial, diagonal, and off-axial/off-diagonoal elements. The exact solution to the CDA for Maxwell's equations in (29) - (32) allows individual contributions from each set of contributing dipoles to be distinguished. Figure 2 shows real and imaginary components of contributions from axial (S_I,(30)), diagonal (S_{II},(31)), and off-axial off-diagonal (S_{III},(32)) terms to the retarded dipole sum. The y-axis scale from -100 to +100 and the x-axis scale from 0 to 4.5 illustrate the behavior of the terms over an extended range of A/π to show the periodic behavior of the divergences. The S_{II} and S_{III} components introduce a divergence at $A = \sqrt{2}/\pi$. When added together, as in (29), the sum of these contributions produces a singularity in S at $A = \sqrt{2}/\pi$ (as in Figure 5). Additional divergences observed to occur in S_{II} and S_{III} components at integer multiples of $A = \sqrt{2}/\pi$ do not produce such a singularity. Divergences occur in the S_I term at integer multiples of $A = 2\pi$ as indicated by (35). Divergences in the S_{III} term also appear at multiples of $A = 2\pi$, although the source of these is not transparent in (32). Adding contributions from all points in the 2D array smooths the periodic behavior of the complete retarded dipole sum.

Real and imaginary solution components: near-, induction- and far-field contributions. Figure 3 shows real and imaginary components of contributions from static (r^{-1}, SZ), induction (r^{-2},IZ), and far-field (r^{-3},RZ) radiative contributions to the retarded dipole sum that arise in S_I, S_{II}, and S_{III}. The magnitude of the near-field contribution in the range shown is less than 5. In contrast to assumptions made by previous exact solutions which neglected effects of far-field radiation, the magnitude of radiative contribution is clearly non-negligible: it dwarfs the intermediate and near-field contributions. Except for the singularity at $A = \sqrt{2}\pi$, the same divergences occur in Figure 3 as were seen in Figure 2. Grouping the contributions into static, intermediate, and far-field terms simplifies illustration of the retarded dipole sum.

Figure 4 shows real and imaginary components of the complete retarded dipole sum as defined in (29). The large divergences that occur periodically as shown in (35) and (36) arise primarily due to the contributions from far-field radiation. As grid size becomes very large the magnitude of these divergences becomes infinite. The behavior of the retarded dipole sum is a strong function of grid size near the divergence points. Larger grid sizes approximate behavior near these points more accurately.

Convergence of the exact solution. To confirm the exact solution, Figure 5 compares real and imaginary parts of S for an infinite periodic lattice, given by (29), with values of S computed for a center particle in a finite array, in the array size range $L = [5, 40]$, by substituting (17) into (27). As L increases, peak values of the computed real part red shift and narrow until at $L > 20$ (400 nanoparticles), the finite S begins to exhibit a narrow lineshape similar to the exact solution for an infinite array. For $L = 100$ the peakwidth of the finite S is essentially equal to the infinite solution; only its amplitude remains smaller. The vertical displacement in Figure 5 between the finite S sum and the infinite solution in (29) arises from constituent sinusoidal functions of the infinite solution, that are not present in (27) after performing the rectilinear transformation on the function call $f_{i,j}(k)$ in (16). The exact solution is computed using a smaller grid size (400) than in Figure 4 (1000), which increases the oscillation in the computed values. Inset in Figure 6 are the times (in seconds) required to compute values of the finite retarded dipole sum at integer values of λ as L increases. Computation was performed using an AMD Turion dual-core CPU and computation time was recorded by onboard time analysis routines within the MATLAB^{TM} software package. Computation time on the order of 60 hours is required to compute S^{LxL} for an 80 x 80 array of nanoparticles for 600 distinct wavelengths. This compares with ca. 60 seconds required to compute the infinite solution.

Convergence of computed finite (i.e $L < \infty$) retarded dipole sum values S^{LxL} toward the infinite solution $S(A,D)$ obtained by allowing $N \rightarrow \infty$ in (27) is a semi-inductive method for validating the accuracy of the development in (49)-(68), which provides sufficient accuracy at reasonable computational expense. Computational methods such as FDTD and DDA which approximate (29) for one-dimensional arrays similary show that $S^{L\times L}$ approaches the 1-D analog of (29) as $N \rightarrow \infty$ [4]. However, the computational complexity of these methods is commonly greater than that of calculating $S^{L\times L}$ using (27). To illustrate, computation times on the order of 10 hours per wavelength were reported to estimate the retarded dipole sum for an array of 1000 nanoparticles [3].

Validation of approximate solutions. The exact solution permits validating results from approximations to the retarded dipole sum obtained by finite difference time domain (FDTD) or discrete dipole approaches. As examples, prior computation [38] suggested that divergences disappear as $R < 30nm$ and a new divergence appears around $A = \pi$. This is clarified by the exact solution, which shows divergences persist (albeit at small magnitudes) at all values of R, and the apparent divergence at $A = \sqrt{2}\pi \simeq 4.44$ disappears into a singularity as the grid size increases. Computations by DDA, T-matrix, FDTD require search algorithms to identify locations of the divergences, while the exact solution provides the precise location of divergences in the retarded dipole sum in (35). Beyond results from FDTD and DDA, the exact solution indicates there are an infinite number of

divergences for an infinite array. Validation of (29) using DDA is precluded by inaccuracies reported to arise when interparticle spacing D is equal to incident electromagnetic wavelength λ [8].

Application of the exact solution. The exact solution in (29) - (32) shows the value of λ at which divergences occur is independent of other system parameters including size and dielectric constant of the dipolar inclusion, and dielectric constant of the surrounding medium. This specificity, and rapid calculation, make application of (29) preferable for multivariate characterization or optimization, relative to either evaluating large arrays like (27) using the finite retarded dipole sum, or using approximations like FDTD. For example, $S(A, D)$ may be employed to determine cross sections for absorption, scattering, and extinction due to dipole-dipole interactions of inclusions in a homogenous host, so-called extended Maxwell-Garnett composites in which inclusions are not small compared to wavelength described by effective-medium theory [39]. The exact solution to the CDA is being compared with results from T-matrix, CDA computations, and experiments to assess validity of CDA for describing arrays of NP with $R > 50nm$ in which multipolar dynamic depolarization and radiative damping may decrease scattering. The exact solution enables rapid, precise estimations of the effect of changes in λ and D on optical cross sections, across large ranges in values of these parameters. With the exact solution in hand, values of k and D may be selected to produce divergences at a desired location (e.g. a semiconductor bandgap).

Section 4: Conclusion

An exact solution to the coupled dipole approximation of Maxwell's equations for electromagnetic interactions in periodic dipole arrays that include contributions due to far-field radiation has been derived. This solution precisely identifies the source, location, and magnitude of periodic divergences in real and imaginary components of the retarded dipole sum which arise solely due to the wavenumber and grid size independent of particle size and dielectric constants of particle and environment. Contributions to the retarded dipole sum from near-field, intermediate, and far-field radiative interactions are distinguished. Far-field radiative contributions are significant, outweighing near-field and intermediate effects. Computed values of real and imaginary components of the retarded dipole sum in finite arrays converge toward the exact solution as array size increases.

Deriving the new solution produces a new method to optimize design of periodic subwavelength nanostructures for a wide range of applications. The improved computational speed of the exact solution allows for multivariate optimization of NP arrays. A single variable interpolation technique has been introduced which allows optimizing design of a sensor using a single retarded dipole array of sufficient grid size. The new solution allows straightforward implementation and design of NP arrays for a wide range of applications. It is a useful benchmark for analyzing performance of computational analysis. Phenomena observed in physical subwavelength nanostructures of increasing complexity can be interpreted with the solution. Divergences identified by the solution correspond to extraordinary spectral resonances in arrays of holes or particles. Applications to improve surface enhanced spectroscopies, sensors, integrated circuit waveguides, optical filters, photovoltaics and antennae are envisioned.

ACKNOWLEDGEMENTS. This work was supported in part by NSF CMMI-0909749. The authors would like to acknowledge valuable conversations with Phillip Blake and Wonmi Ahn.

AUTHOR CONTRIBUTIONS. D.K.R. organized and directed the work and the preparation of the manuscript. B.J.T. performed computational derivation, analysis, and prepared drafts of the model section and first three paragraphs of the discussion section. B.M.H. performed computations and analysis to assist in revising the manuscript and prepared drafts of the remaining paragraphs of the discussion section..

1. **Meier M, Liao PF, Wokaun A (1985) Enhanced Fields on rough surfaces: dipolar interactions among particles of sizes exceeding the Rayleigh limit.** *J Opt Soc Am B* **2:931-949.**
2. **Markel VA (1993) Coupled-dipole approach to scattering of light from a one-dimensional periodic dipole structure.** *J Mod Opt* **40:2281-2291.**
3. **Zhao LL, Schatz GC, Kelly L (2003) The extinction spectra of silver nanoparticle arrays: inuence of array structure on plasmon resonance wavelength and widths.** *J Phys Chem B* **107:7343-7350.**
4. **Zou S, Schatz GC, Janel N (2004) Silver nanoparticle array structures that produce remarkably narrow plasmon lineshapes.** *J Chem Phys* **120:10871-10875.**
5. **Erin M et al. (2004) Investigating narrow plasmons in nanoparticle arrays fabricated using electron beam lithography.** *SPIE Proceedings* **5513:22-29.**
6. **Zou S, Schatz GC (2004) Narrow plasmonic/photonic extinction and scattering line shapes for one and two dimensional silver nanoparticle arrays.***J Phys Chem* **121:12606-12612.**
7. **Fung KH, Chan CT (2007) Plasmonic modes in periodic metal nanoparticle chains: a direct dynamic eigenmode analysis.** *Opt. Lett.* **32:973-975**
8. **Yin S, Luo X, Du C, Zhang Y, Deng Q (2008) The coupled electric field effects on localized surface plasmon resonance in nanoparticle arrays.** *J Appl Phys* **104:024308.**
9. **Lamprecht B, Lechner RT, Ditlbacher H, Krenn JR, Leitner A, Aussenegg FR, Schider G (2000) Metal Nanoparticle Gratings: Inuence of Dipolar Particle Interaction on the Plasmon Resonance.** *Phys Rev Lett* **84:4721-4725.**
10. **Haynes CL, Zhao LL, Van Duyne RP, Schatz GC, Gunnarsson L, Prikulis J, Kasemo B Kall M, McFarland Ad (2003) Nanoparticle Optics: The Importance of Radiative Dipole Coupling in Two-Dimensional Nanoparticle Arrays.** *J Phys Chem B* **107:7337-7342.**
11. **Chamberlin DR, Sultana KA, Chow EK, Sigalas MM, Liu M, Grot AC, Fan S, Wang Z (2005) SERS and plasmon resonance of engineered nanoparticle arrays.** *Preceedings of SPIE* **592708:1-8.**
12. **Felidj N, Aubard J, Levi G, Hohenau A, Krenn JR, Aussenegg FR, Laurent G (2005) Grating-induced plasmon mode in gold nanoparticle arrays.** *J Chem Phys* **123:221103.**
13. **Hicks EM, Gunnarsson L, Zou S, Kasemo B, Kall M, Spears KG, Schatz GC, Van Duyne RP, Rinpnas.bstdzevicius T (2005) Controlling plasmon line shapes through diractive coupling in linear arrays of cylindrical nanoparticles fabricated by electron beam lithography.** *Nano Letters* **5:1065-1070.**
14. **Chumanov G, Malynych SZ (2006) Nanoparticle Assemblies and Superstructures** *In Nanoparticle Assemblies, Superstructures* **pp. CRC press Boca Raton, FL.**
15. **McLellan E, Rindzevicius T, Kall M, Zou S, Spears K, Schatz G, Van Duyne R, Gunnarsson L (2005) Plasmonic and Diffractive Coupling in 2D Arrays of Nanoparticles produced by Electron Beam Lithography.** *Mater Res Soc Symp Proc* **951:E09-20.**
16. **Govyadinov AA, Markel VA (2008) From slow to superluminal propagation: Dispersive properties of surface plasmon polaritons in linear chains of metallic nanospheroids.** *Phys Rev B* **78:035403.**
17. **Carron KT, Meier M, Wokaun A, Fluhr W (1986) Resonances of two-dimensional particle gratings in surface-enhanced Raman scattering.** *J Opt Soc Am B* **3, 430-440.**
18. **Felidj N, Levi G, Krenn JR, Salerno M, Schider G, Lamprecht B, Leitner A, Aussenegg FR, Aubard J (2002) Controlling the optical response of regular arrays of gold particles for surface-enhanced Raman scattering.** *Phys Rev B* **65:075419.**
19. **Genov DA, Shalaev VM, Wei A, Sarychev AK (2004) Resonant Field Enhancements from Metal Nanoparticle Arrays.** *Nano Letters* **4, 153-158.**
20. **Li H, Rothberg L (2004) Colorimetric detection of DNA sequences based on electrostatic interactions with unmodied gold nanoparticles.** *PNAS* **101:14036-14039.**
21. **Souza GR, Staquicini FI, Ozawa MG, Snyder EY, Sidman RL, Miller JH, Arap W, Pasqualini R, Christianson DR (2005) Networks of gold nanoparticles and bacteriophage as biological sensors and cell-targeting agents.** *PNAS* **103:1215-1220.**
22. **Zhang Y, Previte MJ, Geddes CD, Aslan K (2008) Plasmonic engineering of singlet oxygen generation.** *PNAS* **105:1798-1802.**
23. **Agrawal A, Wang GD, Wang MD, Nie S, Deo R (2008) Nanometer-scale mapping and single-molecule detection with color-coded nanoparticle probes.** *PNAS* **105, 3298-3303.**
24. **Alu A, Engheta N (2006) Theory of linear chains of metamaterial/plasmonic particles as subdiraction optical nanotransmission lines.** *Phys Rev B* **74:205436.**
25. **Driscoll T, Bason DN, Palit S, Cho SY, Jokerst NM, Smith DR (2007) Tuned permeability in terahertz split-ring resonators for devices and sensors.** *Applied Physics Letters* **91:062511.**
26. **Shegai T, Dadosh T, Zhang Z, Xu H, Haran G, Li Z (2008) Managing light polarization via plasmon-molecule interactions within an asymmetric metal nanoparticle trimer.** *PNAS* **105, 16448-16453.**
27. **Jackson JB, Halas NJ (2004) Surface-enhanced Raman scattering on tunable plasmonic nanoparticle substrates.** *PNAS* **101:17930-17935.**
28. **GC, Schatz (2007) Using theory and computation to model nanoscale properties.** *PNAS* **104:6885-6892.**
29. **Purcell EM, Pennpacker CR (1973) Scattering and absorption of light by nonspherical dielectric grains.** *Astrohphys J* **186:705-714.**
30. **Yurkin MA, Hoekstra AG (2007) The discrete dipole approximation for simulation of light scattering by particles much larger than the wavelength.** *JQSRT* **106:546-557.**
31. **Yurkin MA, Hoekstra AG (2007) The discrete dipole approximation: an overview and recent developments.** *JQSRT* **106:558-559.**
32. **Merchiers O, Gonzalez F, Saiz JM, Moreno F (2007) Light scattering by an ensemble of interacting dipolar particles with both electric and magnetic polarizabilities.** *Phys Rev A* **76, 043834.**
33. **Khlebstov B, Zharov B, Khlebstov N, Melnikov A (2006) Absorption and scattering of light by a dimer of metal nanospheres: comparison of dipole and multipole approaches.** *Nanotech* **17, 1437-1445.**
34. **Biring S, Wang J, Wang Y, Wang H (2008) Light scattering from 2D arrays of monodispersed Ag-nanoparticles separated by tunable nano-gaps: spectral evolution and analytical analysis of plasmonic coupling.** *Optics Express* **16:15312.**
35. **Klajn R, Grzybowski BA, Bishop KJ (2007) Light-controlled self-assembly of reversible and irreversible nanoparticle suprastructures.** *PNAS* **104:10305-10309.**
36. **Ruppin R (2000) Evaluation of extended Maxwell-Garnett theories.** *Optics Communication* **182:273-279.**
37. **Jackson J (1999)** *Classical Electrodynamics***. (John Wiley Sons, INC.), Third edition, pp. 243-245.**
38. **Zou S, Schatz GC (2006) Theoretical studies of plasmon resonances in one-dimensional nanoparticle chains: narrow lineshapes with tunable widths.** *Nanotechnology* **17:2813-2820.**
39. **Roper D.K, Ahn W, Taylor B, Y. D'Asen Enhanced spectral sensing by electromagnetic coupling with localized surface plasmons on subwavelength structures** *IEEE Sensors* **Accepted Feb. 17, 2009**

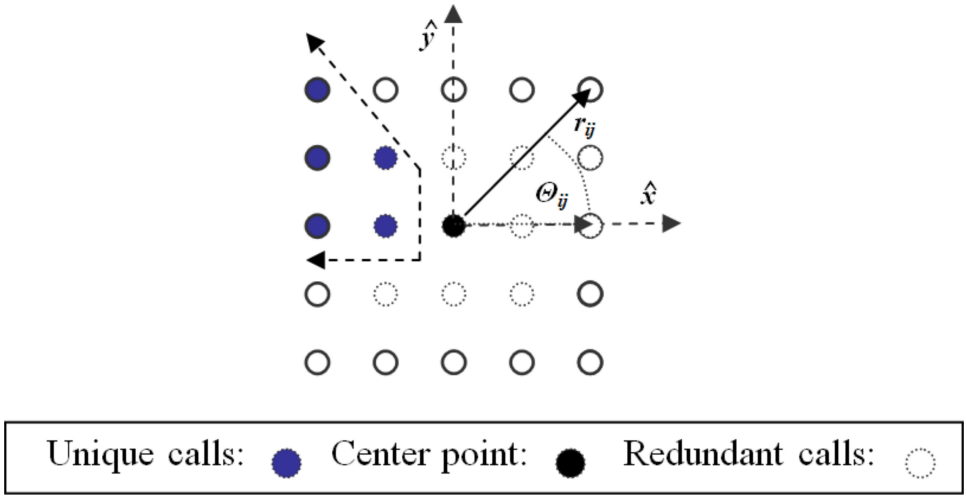

Figure 2.1. *Submitted to Proceedings of the National Academy of Sciences of the United States of America* Grid layout for a $5x5$ square grid where r_{ij} represents the vector between the center point, i, and the point, j. Dotted and solid circles are shown to represent the nested grid concept used to obtain the analytical solution for the retarded dipole sum, S, for a square grid array.

2.2 Supporting PNAS Figures

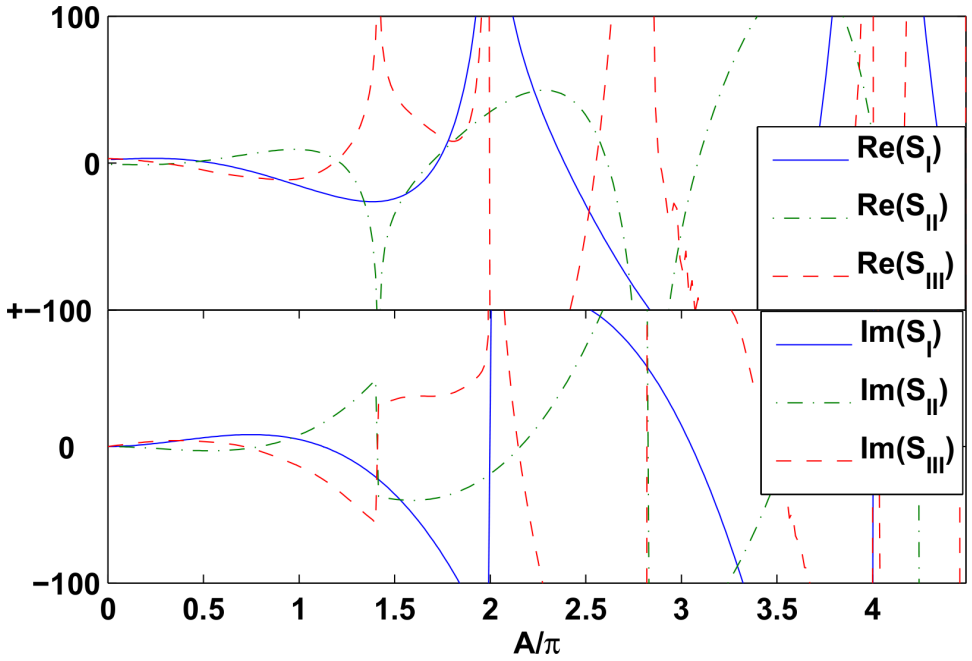

Figure 2.2. *Submitted to Proceedings of the National Academy of Sciences of the United States of America* The retarded dipole sum, S , may also be visualized as the three contributing equations (2.30,2.31,2.32)

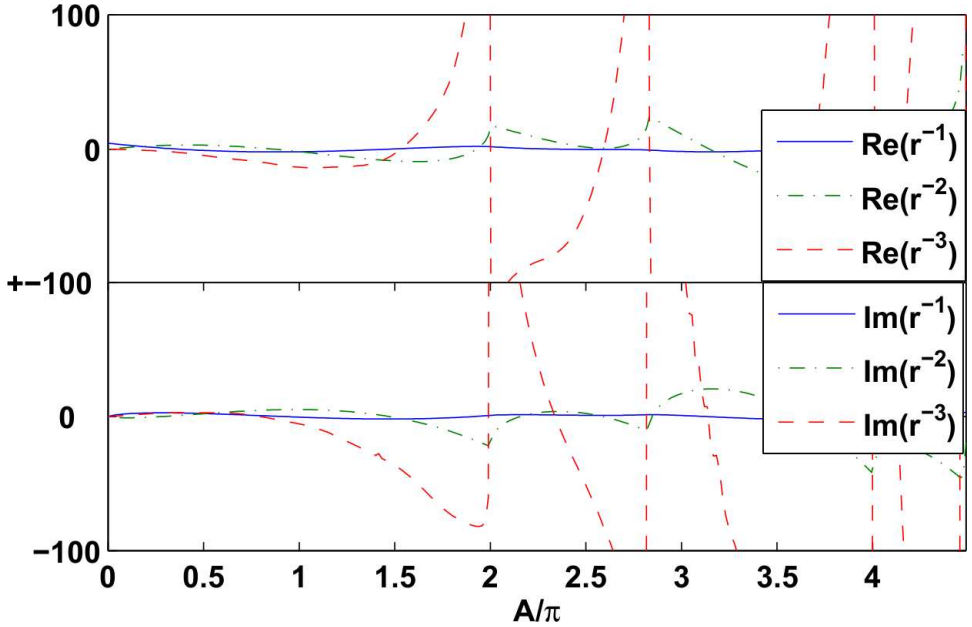

Figure 2.3. *Submitted to Proceedings of the National Academy of Sciences of the United States of America* The retarded dipole sum, S, contribution can be split into static (r^1), intermediate (r^2), and far-field (r^3) contributions for greater understanding

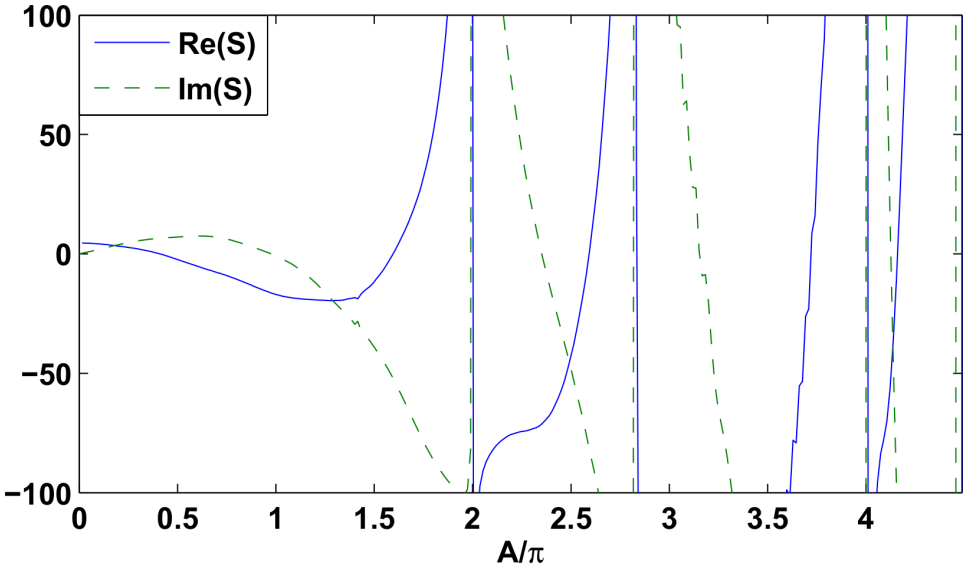

Figure 2.4. *Submitted to Proceedings of the National Academy of Sciences of the United States of America* The singularities in the retarded dipole sum, S, can be seen at the expected locations where $A = \sqrt{2\pi}$ and $A = 2\pi$. An infinite number of singularities exist for an infinite array but due to the D^3 factor in (2.29) singularities resulting in larger A are less significant

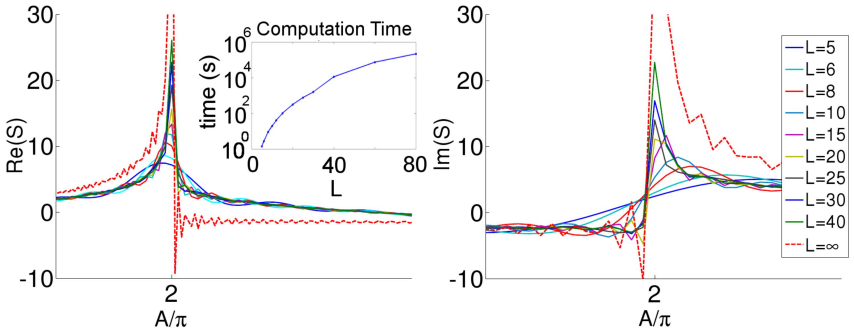

Figure 2.5. *Submitted to Proceedings of the National Academy of Sciences of the United States of America* Computed values of the retarded dipole sum in (27) for finite arrays of size L converge to exact solution for retarded dipole sum in (2.29) for infinite array. As L increases, real and imaginary parts of (2.27) red and blue shift, respectively. For values of $L > 20$, both real and imaginary parts exhibit narrowed lineshapes. The exact solution is computed using a smaller grid size (400) than in Figure 4 (1000), which increases the oscillation in the computed values. Inset: Computation time for finite arrays increases exponentially with L

Supporting Information : Exact solution to the coupled dipole approximation of Maxwell's equations for electromagnetic interactions in periodic nanostructures

Ben J. Taylor[*], **Braden M. Harbin**[†], **and D. Keith Roper**[* †]

[*]University of Utah Department of Chemical Engineering, and [†]University of Arkansas Ralph E. Martin Department of Chemical Engineering

Submitted to Proceedings of the National Academy of Sciences of the United States of America

Appendix

Section A: Electric field derivation

The electric dipole field can be calculated by keeping only the 0th term ($m = 0$) in (6) which gives

$$\mathbf{A}(\mathbf{x}) = \frac{\mu}{4\pi}\frac{\exp(ikr)}{r}\int \mathbf{J}(\mathbf{x}')d^3x'. \quad [37]$$

Integration by parts and combining the relationship, $i\omega\rho = \vec{\nabla}\cdot\mathbf{J}$, from the continuity equation for charge, reduces (37) to

$$\mathbf{A}(\mathbf{x}) = -\frac{\mu}{4\pi}\frac{\exp(ikr)}{r}i\omega\int \mathbf{x}'\rho(\mathbf{x}')d^3x'. \quad [38]$$

From the definition of the electric dipole moment, $\mathbf{p} = \int \mathbf{x}'\rho(\mathbf{x}')d^3x'$, we can further simplify (38) to get

$$\mathbf{A}(\mathbf{r},\mathbf{p}) = -\frac{\mu}{4\pi}\frac{\exp(ikr)}{r}i\omega\mathbf{p}. \quad [39]$$

The electric field can be found by substituing (39) into (5):

$$\begin{aligned}
\mathbf{E}(\mathbf{r},\mathbf{p}) &= \frac{ic}{k}\vec{\nabla}\times(\vec{\nabla}\times\mathbf{A}) & [40]\\
&= \frac{c}{k}\frac{\omega\mu}{4\pi}\left\{\vec{\nabla}\left(\frac{ike^{ikr}}{r}-\frac{e^{ikr}}{r^2}\right)\times(\hat{\mathbf{r}}\times\mathbf{p})+\left(\frac{ike^{ikr}}{r}-\frac{e^{ikr}}{r^2}\right)\vec{\nabla}\times(\hat{\mathbf{r}}\times\mathbf{p})\right\} & [41]\\
&= \frac{1}{4\pi\epsilon}\left\{\left(\frac{2}{r^3}-\frac{2ik}{r^2}-\frac{k^2}{r}\right)\hat{\mathbf{r}}\times(\hat{\mathbf{r}}\times\mathbf{p})+\left(\frac{ik}{r}-\frac{1}{r^2}\right)\vec{\nabla}\times(\hat{\mathbf{r}}\times\mathbf{p})\right\}e^{ikr} & [42]
\end{aligned}$$

where $\hat{\mathbf{r}} = \mathbf{r}/r$ is a unit vector.
The vector identity $\vec{\nabla}\times(\hat{\mathbf{r}}\times\mathbf{p}) = -\mathbf{p}/r - (\mathbf{p}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}}/r$ can be used to further simplify (42) to obtain [1]

$$\mathbf{E}(\mathbf{r},\mathbf{p}) = \frac{1}{4\pi\epsilon}\left\{\left(\frac{2}{r^3}-\frac{2ik}{r^2}-\frac{k^2}{r}\right)\hat{\mathbf{r}}\times(\hat{\mathbf{r}}\times\mathbf{p})-\left(\frac{ik}{r^2}-\frac{1}{r^3}\right)(\mathbf{p}+(\mathbf{p}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}})\right\}e^{ikr} \quad [43]$$

which can be reduced by substituting in Lagrange's formula $\hat{\mathbf{r}}\times(\hat{\mathbf{r}}\times\mathbf{p}) = (\hat{\mathbf{r}}\cdot\mathbf{p})\hat{\mathbf{r}} - (\hat{\mathbf{r}}\cdot\hat{\mathbf{r}})\mathbf{p}$ and factoring like terms to produce

$$\begin{aligned}
\mathbf{E}(\mathbf{r},\mathbf{p}) &= \frac{1}{4\pi\epsilon}\left\{\left(\frac{1}{r^3}-\frac{ik}{r^2}\right)[3(\hat{\mathbf{r}}\cdot\mathbf{p})\hat{\mathbf{r}}-\mathbf{p}]-\frac{k^2}{r}[(\hat{\mathbf{r}}\cdot\mathbf{p})\hat{\mathbf{r}}-\mathbf{p}]\right\}e^{ikr} & [44]\\
&= \frac{1}{4\pi\epsilon}\left\{\left(\frac{1}{r^3}-\frac{ik}{r^2}\right)[3(\hat{\mathbf{r}}\cdot\mathbf{p})\hat{\mathbf{r}}-\mathbf{p}]+\frac{k^2}{r}[(\hat{\mathbf{r}}\times\mathbf{p})\times\hat{\mathbf{r}}]\right\}e^{ikr} & [45]
\end{aligned}$$

which is the general form of the electric field for an oscillating dipole which corresponds to (6).

Reserved for Publication Footnotes

In the $\hat{\mathbf{x}}$-$\hat{\mathbf{y}}$ plane the unit vector has cartesian coordinates $\hat{\mathbf{r}} = \cos(\theta)\hat{\mathbf{x}} + \sin(\theta)\hat{\mathbf{y}}$. If the polarization vector is parallel with $\hat{\mathbf{x}}$ then θ is the angle between the two vectors $\hat{\mathbf{r}}$ and $\mathbf{p}$ and the polarization can be defined as $\mathbf{p} = P_x\hat{\mathbf{x}}$. Substituting the cartesian expression for the polarization and position vector we obtain a two component vector field defined by

$$\mathbf{E_x} = \frac{e^{ikr}}{4\pi\epsilon}\left\{\left(\frac{1}{r^3} - \frac{ik}{r^2}\right)(3\cos^2(\theta) - 1) + \frac{k^2\sin^2(\theta)}{r}\right\}P_x\hat{\mathbf{x}} \quad [46]$$

$$\mathbf{E_y} = \frac{e^{ikr}}{4\pi\epsilon}\left\{\left(\frac{1}{r^3} - \frac{ik}{r^2}\right)3\cos(\theta)\sin(\theta) + \frac{k^2\cos(\theta)\sin(\theta)}{r}\right\}P_x\hat{\mathbf{y}}. \quad [47]$$

In the case of an array of particles symmetric about both $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, the $\mathbf{E_y}$, contributions sum to zero and the contribution from neighboring dipole fields can be written as

$$\sum_{j\neq i}\mathbf{E}\left(\mathbf{r}_{ij}, \mathbf{p}_j\right) = \sum_{j\neq i}\frac{e^{ikr_{ij}}}{4\pi\epsilon}\left\{\left(\frac{1}{r_{ij}^3} - \frac{ik}{r_{ij}^2}\right)(3\cos^2(\theta_{ij}) - 1) + \frac{k^2\sin^2(\theta_{ij})}{r_{ij}}\right\}\mathbf{p_j} \quad [48]$$

as seen in (13).

Section B: Analytical Solution Derivation

Beginning with the final result from the derivation for $S^{L\times L}$ (27)

$$S^{L\times L} = \frac{1}{4\pi\epsilon}\sum_{m=1}^{N}\left[\overbrace{2(f_{m,0} + f_{0,m})}^{axial} + \overbrace{4f_{m,m}}^{diagonal}\right] + \overbrace{\frac{1}{\pi\epsilon}\sum_{m=2}^{N}\sum_{n=1}^{m-1}[f_{m,n} + f_{n,m}]}^{off-axial/diagonal(OAD)} \quad [49]$$

we take the axial term when $N = \infty$

$$S_{axial} = \frac{1}{4\pi\epsilon}\sum_{m=1}^{\infty}\overbrace{2(f_{m,0} + f_{0,m})}^{axial} \quad [50]$$

and substitute in the expression for $f_{x,y}(k, D)$ in (50) and simplify to get

$$S_{axial}(k, D) = \frac{1}{4\pi\epsilon}\sum_{m=1}^{\infty}e^{ikDm}\left\{\overbrace{2\left(\frac{(2) + (-1)}{D^3m^3}\right)}^{Static\ (SZ)} + \overbrace{2\left(\frac{(-i2k) + (ik)}{D^2m^2}\right)}^{Intermediate\ (IZ)} + \overbrace{2\left(\frac{(0) + (k^2)}{Dm}\right)}^{radiation\ (RZ)}\right\} \quad [51]$$

$$= \frac{1}{2\pi\epsilon}\sum_{m=1}^{\infty}e^{ikDm}\left\{\frac{1}{D^3m^3} - \frac{ik}{D^2m^2} + \frac{k^2}{Dm}\right\}. \quad [52]$$

We introduce $A = kD$ which yields

$$S_{axial}(A, D) = \frac{1}{2\pi\epsilon D^3}\left\{\sum_{m=1}^{\infty}\frac{\left(e^{iA}\right)^m}{m^3} - iA\sum_{m=1}^{\infty}\frac{\left(e^{iA}\right)^m}{m^2} + A^2\sum_{m=1}^{\infty}\frac{\left(e^{iA}\right)^m}{m}\right\} \quad [53]$$

$$= \frac{1}{2\pi\epsilon D^3}\left\{\overbrace{P_3\left(e^{iA}\right)}^{SZ} - \overbrace{iAP_2\left(e^{iA}\right)}^{IZ} - \overbrace{A^2\ln\left(1 - e^{iA}\right)}^{RZ}\right\} \quad [54]$$

$$= \frac{1}{2\pi\epsilon D^3}\left\{S_I(A)\right\}. \quad [55]$$

where P_3 and P_2 are polylogarithms defined in (33) and the RZ term contains the anlytical solution to P_1, $P_1 = -ln(1 - x)$.

Likewise, the diagonal term is derived beginning with

$$S_{diagonal} = \frac{1}{4\pi\epsilon}\sum_{m=1}^{\infty}\overbrace{4f_{m,m}}^{diagonal} \quad [56]$$

we substitute (17) and simplify to get

$$S_{diagonal}(k,D) = \frac{1}{\pi\epsilon}\sum_{m=1}^{\infty} e^{ikD\sqrt{2m^2}}\left\{\overbrace{\frac{(2m^2-m^2)}{D^3\sqrt{2m^2}^5}}^{SZ} - \overbrace{\frac{ik(m^2-2m^2)}{D^2\sqrt{2m^2}^4}}^{IZ} + \overbrace{\frac{k^2m^2}{D\sqrt{2m^2}^3}}^{RZ}\right\} \quad [57]$$

$$= \frac{1}{\pi\epsilon}\sum_{m=1}^{\infty} e^{ikD\sqrt{2}m}\left\{\frac{\sqrt{2}}{8D^3m^3} - \frac{ik}{4D^2m^2} + \frac{\sqrt{2}k^2}{4Dm}\right\}. \quad [58]$$

Substituting in $A = kD$ provides

$$S_{diagonal}(A,D) = \frac{1}{\pi\epsilon D^3}\sum_{m=1}^{\infty}\left\{\frac{\sqrt{2}}{8}\frac{\left(e^{iA\sqrt{2}}\right)^m}{m^3} - \frac{iA}{4}\frac{\left(e^{iA\sqrt{2}}\right)^m}{m^2} + \frac{\sqrt{2}A^2}{4}\frac{\left(e^{iA\sqrt{2}}\right)^m}{m}\right\} \quad [59]$$

$$= \frac{1}{4\pi\epsilon D^3}\left\{\overbrace{\frac{\sqrt{2}}{2}P_3\left(e^{iA\sqrt{2}}\right)}^{SZ} - \overbrace{iAP_2\left(e^{iA\sqrt{2}}\right)}^{IZ} - \overbrace{\sqrt{2}A^2\ln\left(1-e^{iA\sqrt{2}}\right)}^{RZ}\right\} \quad [60]$$

$$= \frac{1}{4\pi\epsilon D^3}\left\{S_{II}(A)\right\}. \quad [61]$$

The remaining off-axial and off-diagonal terms can be derived beginning with

$$S_{OAD} = \overbrace{\frac{1}{\pi\epsilon}\sum_{m=2}^{N}\sum_{n=1}^{m-1}[f_{m,n}+f_{n,m}]}^{off-axial/diagonal(OAD)} \quad [62]$$

we substitute (17) to produce

$$S_{OAD}(k,D) = \frac{1}{\pi\epsilon}\sum_{m=2}^{\infty}\sum_{n=1}^{m-1}\left\{\overbrace{\frac{e^{ikD\sqrt{m^2+n^2}}}{D^3\sqrt{m^2+n^2}^3}}^{SZ} - \overbrace{\frac{ike^{ikD\sqrt{m^2+n^2}}}{D^2\sqrt{m^2+n^2}^2}}^{IZ} + \overbrace{\frac{k^2e^{ikD\sqrt{m^2+n^2}}}{D\sqrt{m^2+n^2}}}^{RZ}\right\}. \quad [63]$$

Substituting in $A = kD$ provides

$$S_{OAD}(k,D) = \frac{1}{\pi\epsilon D^3}\sum_{m=2}^{\infty}\sum_{n=1}^{m-1}\left\{\overbrace{\frac{(e^{iA})^{\sqrt{m^2+n^2}}}{\sqrt{m^2+n^2}^3}}^{SZ} - \overbrace{iA\frac{(e^{iA})^{\sqrt{m^2+n^2}}}{m^2+n^2}}^{IZ} + \overbrace{A^2\frac{(e^{iA})^{\sqrt{m^2+n^2}}}{\sqrt{m^2+n^2}}}^{RZ}\right\} \quad [64]$$

$$= \frac{1}{\pi\epsilon D^3}\left\{\overbrace{Q_3(e^{iA})}^{SZ} - \overbrace{iAQ_2(e^{iA})}^{IZ} + \overbrace{A^2Q_1(e^{iA})}^{RZ}\right\} \quad [65]$$

$$= \frac{1}{\pi\epsilon D^3}\left\{S_{III}(A)\right\} \quad [66]$$

where $Q_z(x)$ is a double sum defined as

$$Q_z(x) = \sum_{m=2}^{\infty}\sum_{n=1}^{m-1}\frac{x^{\sqrt{m^2+n^2}}}{(m^2+n^2)^{z/2}}. \quad [67]$$

We combine results from (55, 61, 66) to obtain the final expression for S

$$S(A,D) = \frac{1}{4\pi\epsilon D^3}(2S_I(A) + S_{II}(A) + 4S_{III}(A)) \quad [68]$$

which is the final form of the new analytical solution beginning in (29).

Section C: Evaluation of off-axial and off-diagonal terms in S

In every continuous span of (32)[1], i.e. A=[0,$\sqrt{2}\pi$],[$\sqrt{2}\pi$,2π],..., a finite number of interpolation points may be used to fit (32) using piecewise cubic interpolation. Factoring out $1/D^3$ in (29) has allowed (32) to be represented as a function of A alone, which allows one dimensional interpolation as opposed to bicubic interpolation in variables k and D. To optimize the spacing between the finite number of points, a Monte Carlo simulation was used. The Monte Carlo simulation was chosen due to its simplicity to code with respect to other optimization routines, such as the genetic algorithm or game theory. The Monte Carlo simulation was able to reproduce identical results when sufficiently large simulations were run.

Table 1 identifies interpolation points obtained across the interval $A = [0, 2\pi]$. Over this interval, a value of $R^2 \geq 0.99$ was obtained using piecewise cubic interpolation. The number and/or spacing of interpolation points in the Monte Carlo simulation could be increased to obtain higher accuracy. New interpolation points are required to extend computation of the exact solution to values $A/\pi > 4.2$ (e.g. into the infrared region). Figure 1 shows the interpolation points together with the behavior of real and imaginary components of (32) – off-axis and off-diagonal contributions to the retarded dipole sum – plotted versus A/π up to a value just larger than 2. The simulation shown used 1000^2 NPs. Oscillations observed in Figure 2 dampen out with larger grid sizes. The divergence at $A/\pi = \sqrt{2}$ becomes a singularity when contributions from (30) and (31) are included to obtain the complete retarded dipole sum (shown in Figure 5), whereas the divergence at $A/\pi = 2$ remains a divergence. The y-axis scale from -200 to +200 and the x-axis scale from 0 to 2.1 were selected to show the concentration of the interpolation points near the divergences.

It is useful to note that the retarded dipole sum need not be recalculated when another variable besides k and D is adusted during multivariable optimization on any mathematical derivative of the retarded dipole, like polarization, optical cross section, or electric field in (15). As an example, when examining effects of varying particle diameter or dielectric constant on optical cross section, only a single set of calculated or interpolated values of the retarded dipole sum for a given matrix of k and D is required.

[1] Equations (1) - (36) refer to equations within the main document.

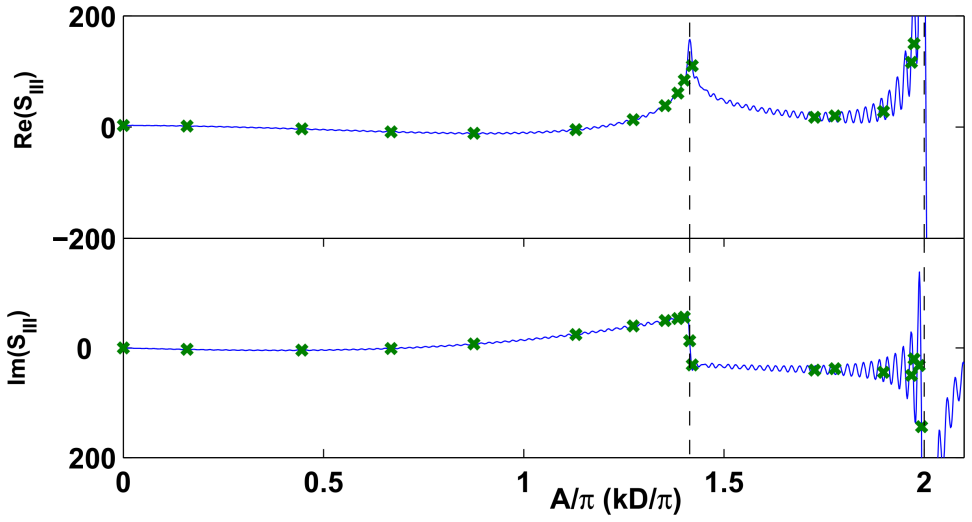

Figure 2.6. *Submitted to Proceedings of the National Academy of Sciences of the United States of America* The behavior of the S_{III} double sum is shown with the interpolation points. This simulation used 1000^2 NPs which produces oscillation

Table 1: Points of interpolation for computing real and imaginary components of (32) using one dimensional piecewise cubic interpolation in A. Interpolation points were determined by using a Monte Carlo simulation to determine optimal spacing. The values reported are the result of a $100,000^2$ NP simulation.

$A = kD$	$Re(S_{III})$	$Im(S_{III})$
0.00	0.3150	0
0.50	-0.1267	0.3467
1.40	-0.4922	-0.4190
2.10	0.2066	-0.8707
2.75	1.4565	-0.5822
3.55	2.0348	2.7089
4.00	-0.5870	5.9506
4.25	-6.0960	8.6120
4.35	-11.3767	9.8267
4.40	-16.9787	10.6193
$\sqrt{2}\pi$	$-\infty(Grid \to \infty)$	10.6193
R^2	0.9981	0.9997
$\sqrt{2}\pi$	$-\infty$	0
4.463	-23.09	-10.67
5.423	-13.85	-10.95
5.583	-4.34	-10.86
5.963	27.45	-2.93

2.3 Design Optimization

To illustrate the usefulness of the new analytical solution, an Ag NP array sensor application will be optimized for sensitivity. The NP array sensor is most sensitive when sharp spectral peaks exist where slight shifts in local refractive index cause dramatic changes in the peak wavelength, as seen in Figure 2.7. With the new model, the existence of extraordinary peaks can be predicted with the knowledge of the location of singularities *prior to computation*. Knowing the precise location of the singularities limits the time needed to locate them as was previously done using brute force search methods [22].

The sensitivity, which is the maximum change in extinction efficiency, can be seen in Figure 2.7 for various interparticle spacings. For an infinite spacing, or single particle spectra, the sensitivity is the lowest because a change in refractive index will have a minimal change in extinction efficiency. However, for the blue curve which is an interparticle spacing of 480nm the peak is very sharp and therefore a small change in refractive index will lead to a large change in extinction efficiency. So a spacing of 480nm would be preferred for

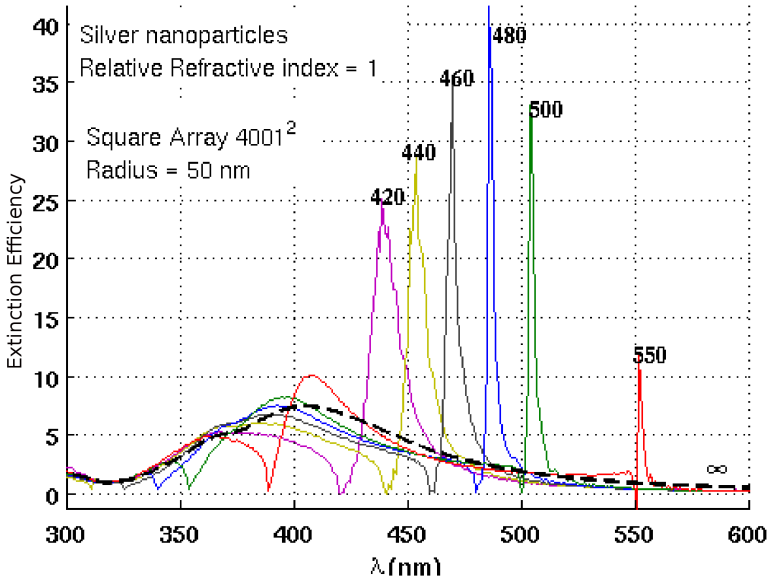

Figure 2.7. *(Table and caption reproduced with permission from IEEE)* Near-field plasmon features (<400nm) with far-field extinction peaks (>450 nm) for 50 nm radius Ag NP array embedded in air. The different colored lines represent different interparticle spacings and the black line is spectra for an infinite spacing. Infinite spacing can also be thought of as a single particle spectra.

sensitivity for silver NP radius of 50nm in a medium with a local refractive index of 1.

Figure 2.8 shows a continuation of an Ag sensor design optimization similar to Figure 2.7 but with higher step-resolution combined with the optimization of particle size. The colormap is linked to sensitivity which is proportional to the maximum slope of the spectral response and the desired optimization fitness parameter. With the new algorithm, Figure 2.8 was created in less than a day and quickly identifies ideal spacing and particle size for the desired target wavelength. These results are significant because they allow for rapid NP sensor design optimization and therefore maximum sensitivity.

In addition to optimizing the NP size and interparticle spacing, the new model can optimize refractive index and even NP material allowing new, unstudied materials to be noticed as potential replacements for Ag and Au NP arrays. The optimization results from the model can be used to match the manufacturing process and minimize the number of experiments needed to achieve the desired output with the aid of the advanced analysis techniques discussed in §4.3-4.5.

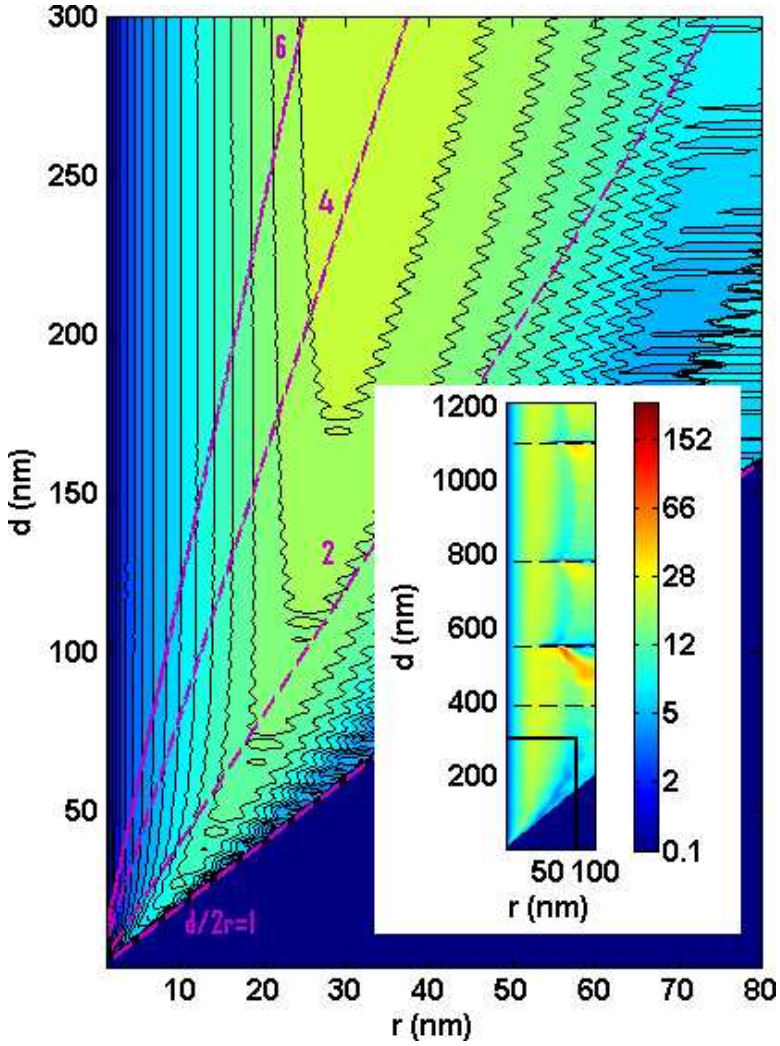

Figure 2.8. *(*Table and caption reproduced with permission from IEEE) Sensitivity of AU NP extinction efficiency (EE) with respect to change in local refractive index, η_o, where sensitivity is inversly proportional to the peak full width half max (FWHM) and defined as the maximum extinction efficiency delta. A nearfield subset from the inset plot shows effects of changing particle radius, r, and interparticle spacing d. The colorscale shows sensitivity values up to 188. Dashed red lines show where relative separation distance, $d/2r$, is constant as r and d are varied. The insert plot shows extraordinary far-field effects predicted by the new algorithm. Dashed lines in the insert plot show discontinuities predicted using the new algorithm with adjacent high sensitivity. This plot allows the search space to be viewed for the optimization problem, and demonstrates the predictability of the singularities. When optimizing this problem in practice for higher dimensions the results are not visualized, but optimized using a metaheuristic global search algorithm.

CHAPTER 3

FABRICATION

3.1 Introduction

Nanoparticle (NP) fabrication is the final step of producing sensors. After electromagnetic theory has been used to optimize the theoretical NP size, spacing, and material, a method must be developed to create such a device. Using advanced electroless deposition (EL), Ahn et al. [2] have developed a low cost technique capable of producing random arrays of Au NPs with varying size and spacings.

3.2 Heating Au Island Films Forms Nanoparticles

Nanoparticles (NP) have been fabricated on Silicon (Si) substrate using various methods including ion implantation and surface etching [16], laser ablation [9], inductively coupled thermal plasmas [11], thermodynamic self organisation [8], and thermal evaporation of thin Au film followed by annealing [6].

Island films and thin films have been formed using electrodeposition, electroless deposition (EL) [15], sputtering [10], and vacuum evaporation [5]. A new NP fabrication technique has been developed which uses EL to create island films and then thermally annealed the island films to create NPs of controllable size distributions. The new EL process has an advantage over other methods for being able to rapidly coat fragile, 3D, or internal surfaces at ambient conditions without requiring expensive specialized equipment or conductive surfaces in the case of electrodepostion. The EL process is also relatively inexpensive for mass fabrication and scales well making it a promising economical solution for NP array applications in the market.

Ahn et al. outlined the EL plating method [2] which involved three steps. To summarize: first, sensitization of quartz slides (GE 124 fused quartz, Chemglass Inc., Vineland, NJ) by immersion in $SnCl_2$ and trifluoroacetic acid followed by rinsing with distilled deonized, degassed (DDD) DDD-H_2O and drying with N_2 gas; second, silver (Ag) deposition was accomplished using $AgNO_3$ followed by rinsing with DDD-H_2O and drying with N_2 gas;

finally, immersion of Ag-derivatized quartz slides into the gold (Au) solution $Na_3Au(SO_3)_2$ for the desired deposition time and rinsing and drying to complete. For a complete recipe detailing concentrations, times, and additional solutions and steps needed for the new EL method along with more experimental results refer to the full paper [2].

Once the Au island films have been created, the transformation from island films to NPs can be seen in Fig 3.1. The first row shows EL island films for (A) 0.33 min (B) 1 min, and (C) 4 min $Na_3Au(SO_3)_2$ deposition times. The second row are the same samples after annealing for three hours at 250 °C. The final step involves heating the annealed samples to 800 °C for 20 min to create spherical NPs, as shown in the bottom row in Fig 3.1.

The transmission and extinction spectra were collected by Ahn [2] who assisted in the collection of SEM images as well in Fig 3.1. Notice as the samples become more spherical

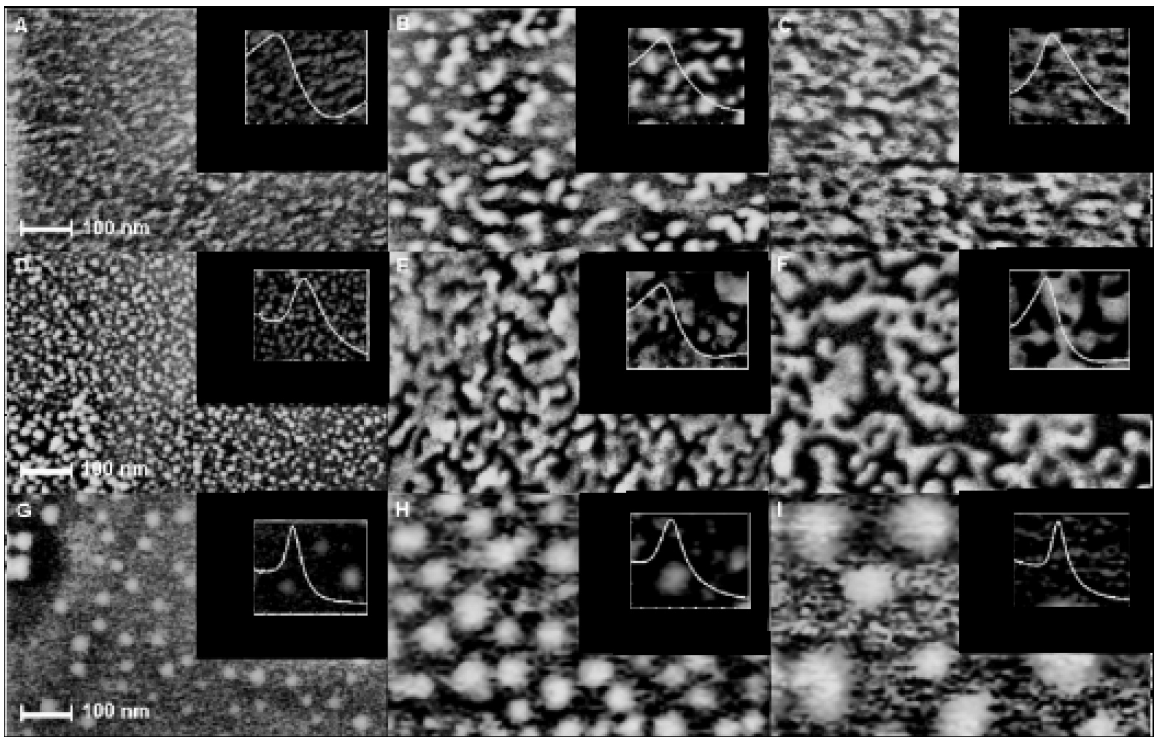

Figure 3.1. (*Figure and caption reproduced with permission from Langmuir* [23]) SEM images of fuzzed quartz slides after immersion (first row) in Au plating solution for (A) 0.33 min, (B) 1 min, and (C) 4 min; after respective immersed slides are annealed (second row, images D, E, and F) for 3 h at 250 °C; and after respective immersed, annealed slides are heated (third row, images G, H, and I) for 20 min at 800 °C. Inset are transmission UV-vis spectra corresponding to each image. Percent transmission is shown in images A-C, E, and F. Extinction in arbitrary units (au) is shown in images D, G-I. All images are the same magnification; see scalebar on the left.

the inset UV-vis spectra show a sharpening resonance peak that is anticipated with NP array theoretical modeling.

Longer deposition times were also studied as well with Fig 3.2 showing (A) 8 min, (B) 30 min, and (C) 60 min $Na_3Au(SO_3)_2$ deposition times. The sharper contrast between Fig 3.1 and Fig 3.2 is a result of the decreased magnification and also a higher quality low vacuum SEM used at Brigham Young University.

The weblike structures seen in Fig 3.2 are NaCl-like crystal structures determined by X-Ray Diffraction (XRD) that were later removed after prolonged rinsing with DDD-H_2O. Manual masking had to be performed to allow data analysis due to the image processing interference created by the crystal structures. The data and statistics from the NP samples in Fig 3.2 and Fig 3.1 are discussed in the following data analysis chapter.

3.3 Multiple Experiments: Drop Method

To reduce the amount of $Na_3Au(SO_3)_2$ needed to plate a thin film and to increase the number of deposition time experiments, an EL drop method was developed with the help of Ahn. After the proposed procedures of Ag and Au deposition on quartz slides, a drop of 1-5 μL $Na_3Au(SO_3)_2$ solution was placed on a slide for various amounts of time. Images (see Fig 3.3) and spectra resulting from drop volumes in this range were similar. To maintain the necessary N_2 to prevent $Na_3Au(SO_3)_2$ from oxidizing, a polypropylene container was used during the deposition period to enclose the sample. Afterward, the Au drop-plated slide was removed and washed with DDD-H_2O and dried with N_2 to create the samples.

The NP diameters size distributions were similar throughout the drop; however, excessive Au deposition was noticed around the droplet perimeter. The excessive Au deposits can

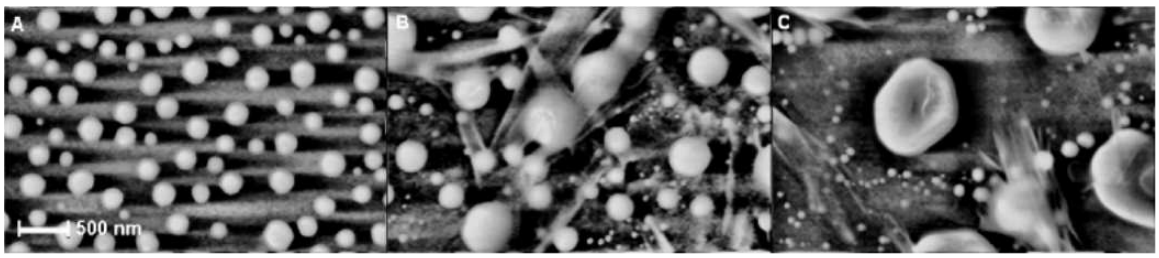

Figure 3.2. (*Figure and caption reproduced with permission from Langmuir* [23]) SEM images of fused quartz slides annealed (250 °C; 3h) *and* heated (800 °C; 20 min) after immersion in Au plating solution for (a) 8 min, (b) 30 min, and (C) 60 min. Weblike strands are NaCl-lie crystal structure removed after rinsing with DDD-H_2O which were identified using X-Ray Diffraction (XRD).

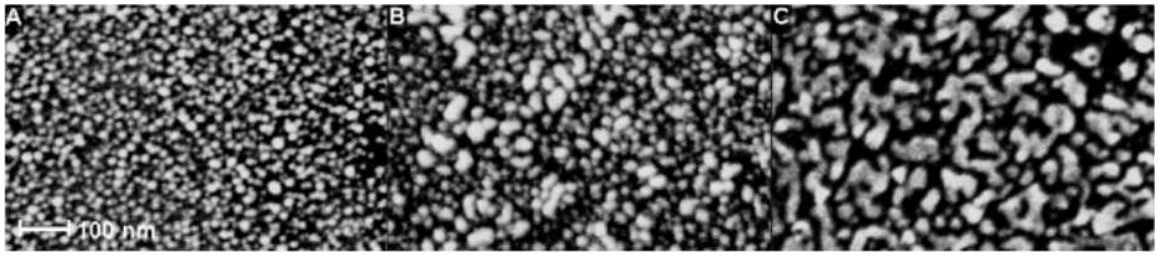

Figure 3.3. (*Figure and caption reproduced with permission from Langmuir* [23]) SEM images after annealing (250 ° C, 3 h) fused quartz slides exposed to microliter droplets of EL plating solution exposed for (A) 0.33 min, (B) 1 min, and (C) 2 min.

be explained by the increased local concentration of $Na_3Au(SO_3)_2$ as the solvent evaporated at the droplet boundary during the deposition time. A low vacuum SEM image can be seen in Fig 3.4 along with an average intensity plot created from Fig 3.4 in Fig 3.5 showing the relative changes in SEM intensity as a function of droplet radius. The charging effects observed near the edge caused varying intensity in the field of view (FOV); therefore, the trending intensity seen towards the edges of the image should not be interpreted as changes in Au concentration.

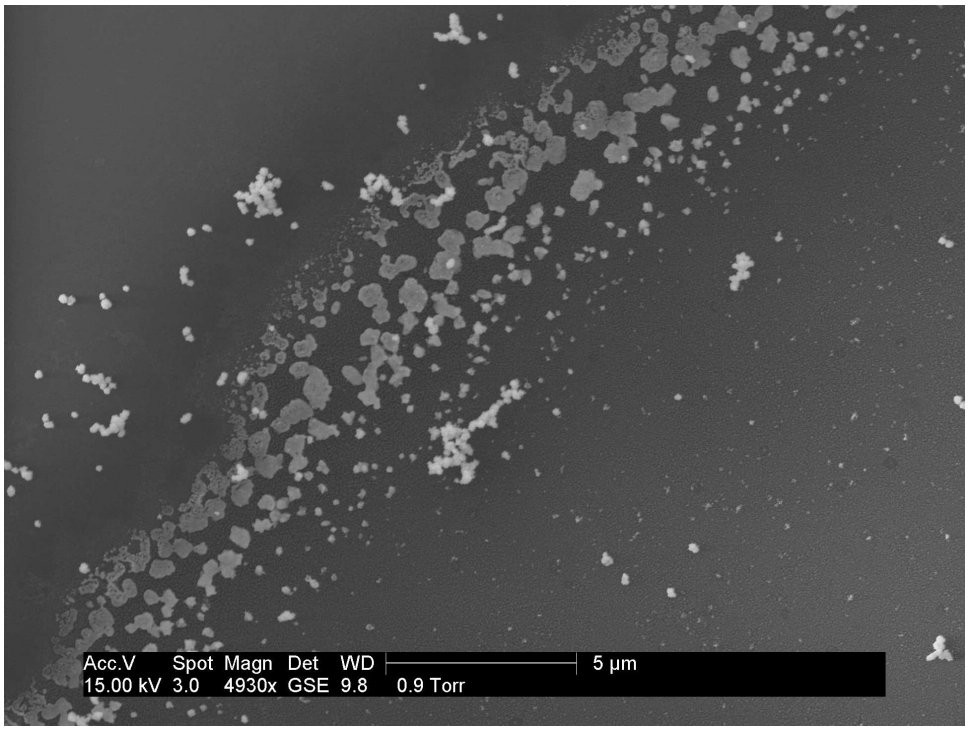

Figure 3.4. Low vacuum SEM image showing increased Au deposits at the drop boundary as the drop evaporated from N_2 convection during the deposition time

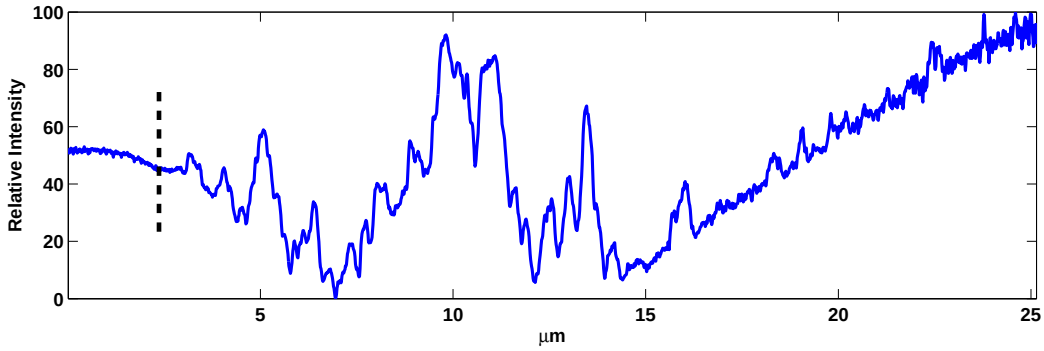

Figure 3.5. Relative pixel intensity from SEM image of droplet edge shows increased Au deposits near the boundary and the change in Au coverage uniformity moving towards the center of the particle. The vertical dashed line represents the best approximation of the outside droplet boundary. Relative slopes at the beginning and end of the plot should not be taken literally since charging may be effecting perimeter intensity in the Field Of View (FOV). *Figure copyright is owned by author*

3.4 Experimental Results Compared to Theoretical

As shown in Fig 3.6, the theoretical results produce a higher extinction efficiency due to the ordered grid. A randomly oriented array of NPs cancels out far-field interactions and therefore would dampen the extinction. The spread between the experimental results is most likely caused by a difference in particle size distributions between the samples which was an artifact from varying spacing. Experimental results reflect relative shifts but not overall fit.

To more accurately test the validity of the theoretical solution, an ordered array with fixed NP size would be required. This could be achieved using lithography, and is the next logical step in this research because theoretically, ordered arrays produce orders of magnitude sensitivity improvements.

3.5 Conclusion

Despite the novelty of the new EL method and its low cost appeal, dramatic sensitivities can only be realized with an evenly spaced array and equal particle sizes. If spacing and particle size are not equal, then the equal polarization assumption (2.15) used in the derivation is not valid. The theoretical work in Section 2.2 shows that far-field effects are additive and therefore exact spacing is preferred to achieve maximum sensitivity.

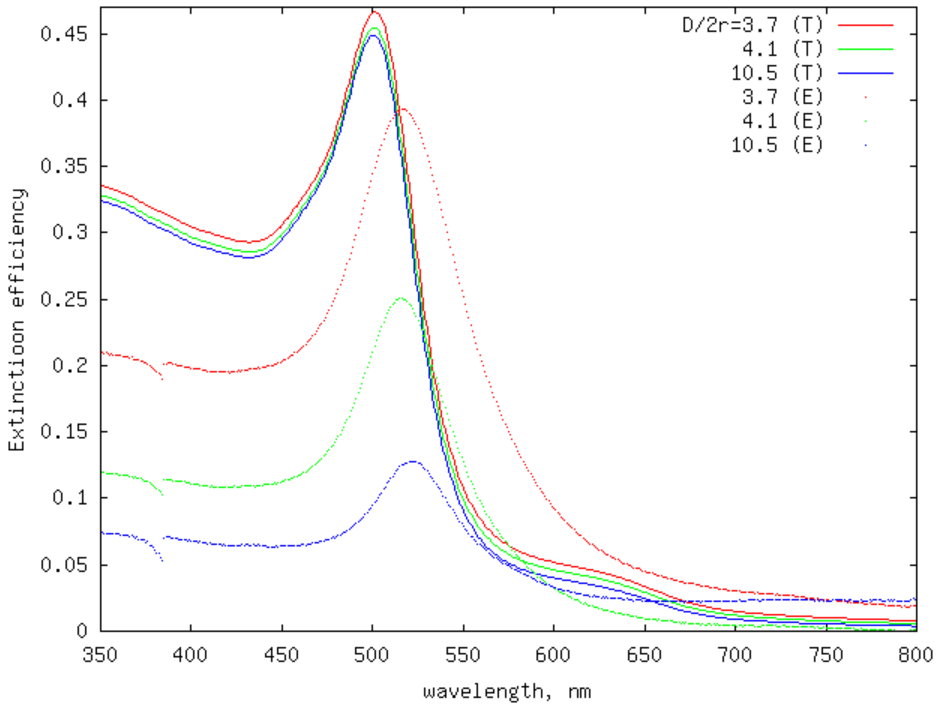

Figure 3.6. Theoretical results deviate from the experimental results for several reasons. The experimental results are created with randomly oriented arrays with a particle size distribution. This violates the equal polarization assumption used in the theoretical derivation (2.15). In order to perform a more rigorous appraisal of the theoretical results, a lithography method of fabrication would be needed to ensure equal spacing and particle size. *Figure copyright is owned by author*

Future work will most likely involve some form of lithography similar to what is used for semiconductor manufacturing. E-beam lithography would be preferred for initial research but mass production of sensors would require a more economical solution such as photolithography. Photolithography methods would require an expensive reticle mask, but would save money long-term.

CHAPTER 4

DATA ACQUISITION AND ANALYSIS

4.1 Data Aquistion

The NP arrays were analyzed using Atomic Force Microscopy (AFM) and a Scanning Electron Microscope (SEM). The SEM images were obtained using a Philips XL30 ESEM FEG (FEI, Hillsboro, Oregon) using a backscattered electron detector mode operating with a 15 kV electron accelerating voltage. The low vacuum setting was used and the pressure was varied independently for each image based on charging and desired resolution over a low vacuum range of 0.4-1.2 Torr. Acquisition preprocessing software automatically stretches to accommodate low- and high-intensity in an image making intensity comparisons between images susceptible to error. The low vacuum setting helped reduce the charging effects that were observed on high vacuum SEM images taken previously which improved the image contrast and accuracy in particle size.

4.2 Watershed-Transform and Shape Analysis

SEM particle analysis is typically performed using threshold algorithms for isolating particles. Fig 4.1 shows that this method can produce inaccurate particle size and shape measurements when adjacent particles are close and when poor image resolution and quality result in clumping. A watershed transformation [13] was used to segment the clumps and provide more accurate size and shape characterization. Briefly, the watershed transform segments the image by relying on the image topology. The original image is transformed into the complement, making local maxima into local minima and vice versa. The watershed transform then begins to fill the now local minima volume, hence the name. While filling, if the local minima volume contacts an adjacent local minima volume, a division is made and the particle becomes segmented.

Two algorithms have been developed to complement the adapted watershed transformation. The first algorithm strips out particles that are touching the border of an image to exclude incomplete particles from the measurements of particle size, shape, and number

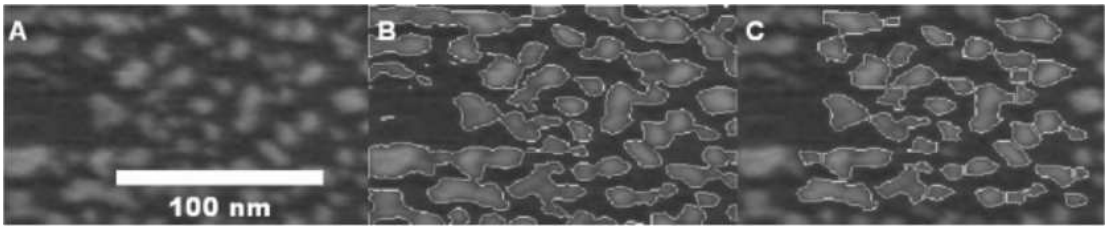

Figure 4.1. (*Figure and caption reproduced with permission from Langmuir* [23]) The watershed method distinguishes particle characteristics more realistically than simple threshold methods when particles are grouped in clusters. (A) Original SEM image of Au-coating on quartz slide. (B) Threshold method of segmentation results in particle "fusion" with size distribution skewed by particles at the border. (C) Watershed method with border removal quantitatively distinguishes particle size, shape, and number density.

density, as seen in Fig 4.1:C. The sodium crystal-like structures in Fig 3.2 that disappeared when samples were rinsed copiously with DDD-H_2O after EL were manually corrected after analysis by the watershed method to differentiate these structures from Au NPs. Crystals selected as particles were manually deselected and missed particles were manually reselected using the open source computer program GIMP (v. 2.2.11; www.gimp.org). The image processing was accomplished using image processing tools (Image Processing Toolbox 6.0) in MATLAB (v. 7.4; MathWorks).

The second algorithm has been developed to analyze particle shape as shown in Fig 4.2. The algorithm selects individual particles and then optimizes the theta, θ, of rotation to minimize the difference in particle shape between the particles. The resulting two-dimensional shapes are then added on top of each other to produce a 3D histogram (Fig 4.3). The resulting 3D histogram can then be used to estimate the probability of finding a particle with a particular shape and easily determine the average sphericity of the Au NPs and other shape-related metrics.

4.3 Applying Watershed to SEM Images

Fig 4.3 shows particle size distributions determined by watershed analysis from SEM images in Fig 3.1 (row 3) and Fig 3.2 for heated Au island films immersed at different times. Fig 4.3:A and Fig 4.3:B show that as immersion time increases from 8 to 30 to 60 min, average particle diameters produced by successive annealing and heating changed from 169.0±2.5 to 266±22 to 97.0±9.5 nm, particle density changed from 7.2 to 4.3 to 12 counts per cm^2 x 10^{-8}, and Full Width Half Max (FWHM) changed from 174 to 60.7 to 113 nm.

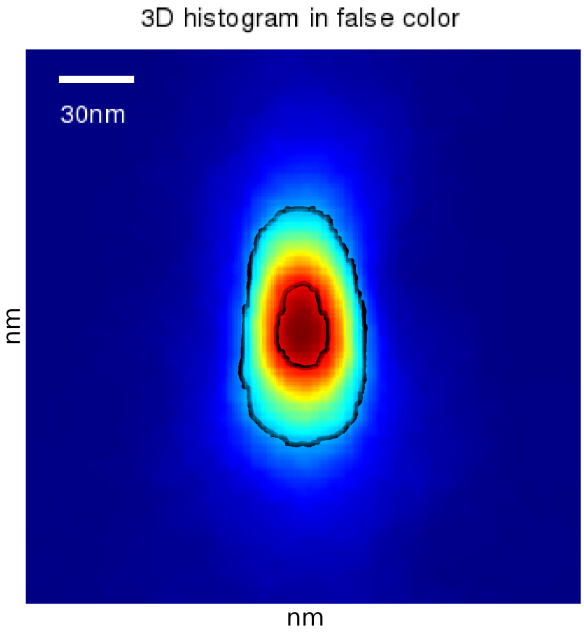

Figure 4.2. Each particle was centered by finding the individual centers of mass and aligned by using the golden search method to maximize the moment about a particular axis. The resulting 3D histogram allows for mean shape determination and shape confidence plots analysis. A shape confidence plot would be an outline shape that contained a quantile of the particle population as shown by the two black contour traces. *Figure copyright is owned by author*

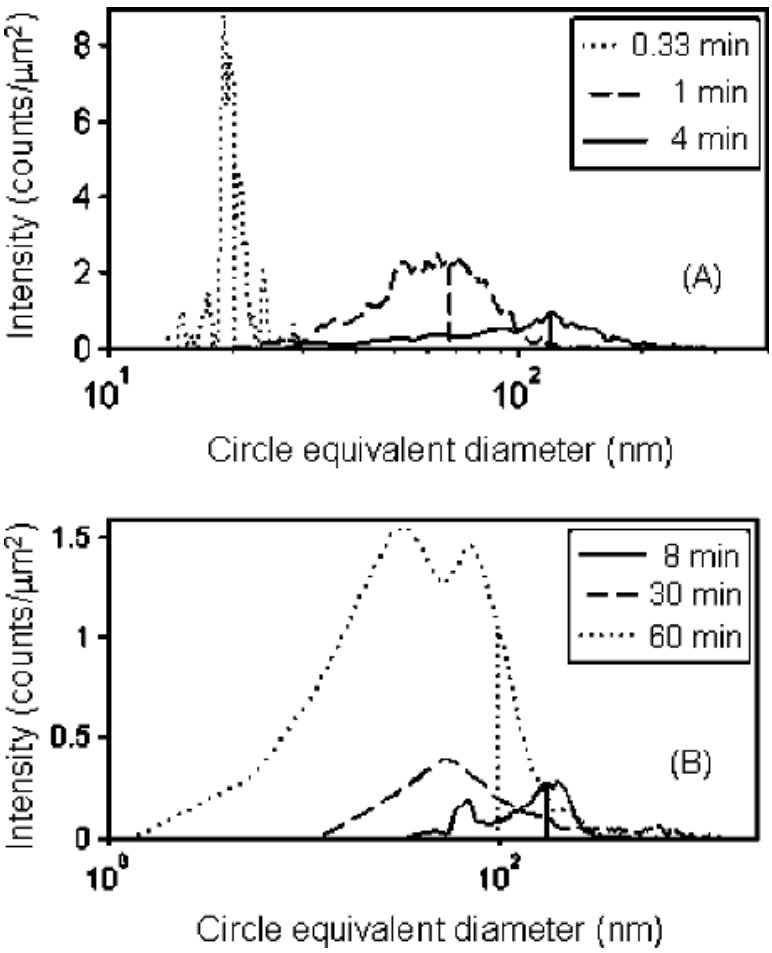

Figure 4.3. (*Figure and caption reproduced with permission from Langmuir* [23]) Mean diameter and density of particles obtained after annealing at 250 °C (3 h) and heating at 800 °C (20 min) shift to larger scales as E.L. deposition time increases from (A) 0.33 to 1 to 4 min. (B) Mean diameter and particle density for annealed and heated films deposited at 8, 30, and 60 min.

Increasing deposition time enlarges Au nanoparticles and broadens the size distribution. EL samples are skewed to smaller particle values by a large number of fine particles identified by watershed analysis in the 30 and 60 min samples.

4.4 Particle Diameters, SEM vs AFM

Table 4.1 shows particle diameters obtained by watershed analysis of SEM images and AFM. The nonlinear behavior of deposition time versus particle size may have been related to the surface area over volume fraction. With a constant deposition flux relative to the surface area, the behavior of deposition would be inversely proportional to the diameter. Therefore, for smaller diameters a faster particle growth rate is expected. AFM diameters exceed SEM diameters in each case except for 4-min deposition, due in part to convolution between probe tip and particle shape. The tip convolution is due to the angle of the tip that prevents the tip of the probe from immediately contacting the substrate once it leaves the particle, therefore leading to an exaggerated particle size. A second computer algorithm has been developed, arbitrarily called "Peak Height Method (PHM)," to determine particle heights from the local minima of particle gradient obtained from inverted AFM images. Table 4.1 shows the mean value of AFM particle heights obtained by this method compared to SEM particle diameters obtained by watershed analysis. As expected, the PHM more closely matches the SEM values than using watershed on AFM data due to the elimination of AFM tip convolution.

Table 4.1. (*Table and caption reproduced with permission from Langmuir [23]*) The AFM (PHM) fits the SEM (WM) data better than the AFM (WM), which overpredicts the particle size, due to tip convolution.

time (min)	particle diameter (nm)		
	SEM(WM)	AFM(WM)	AFM(PHM)
1	67.7±1.1	111.9±3.9	42.3±1.1
4	119.6±4.2	114.1±4.4	49.98±0.82
8	169.4±2.5	241.3±4.6	116.7±2.6
30	266±22	764±80	183±15
60	97±13	580±42	136.9±8.7

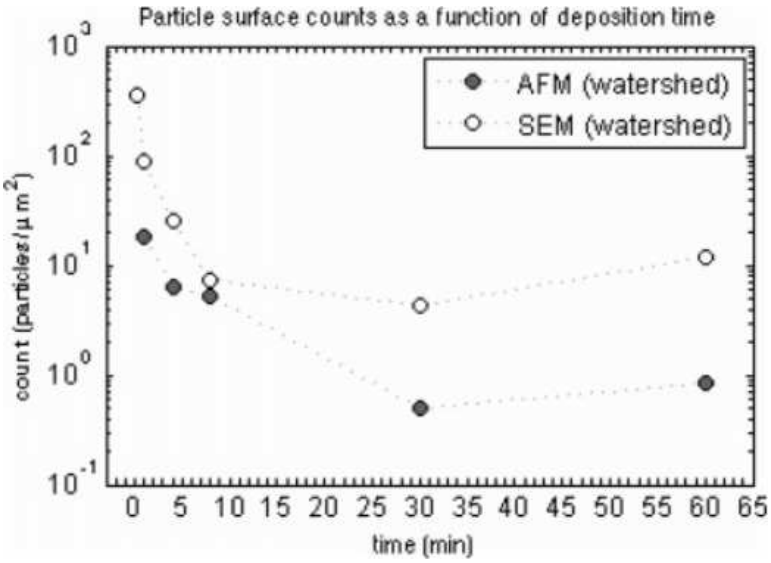

Figure 4.4. (*Figure and caption reproduced with permission from Langmuir* [23]) Particle densities estimated by watershed analysis of AFM (filled circles) and SEM (open circles) images for immersion times at 0.33, 1, 4, 8, 30, and 60 min. The 0.33 min sample was damaged prior to AFM imaging.

4.5 Particle Densities, SEM vs AFM

Fig 4.4 compares particle densities obtained by watershed analysis of AFM and SEM images. To find the particle densities on the Si surface, software counted the number of particles and divided by the total analyzed surface area. Compared with watershed analysis of SEM, AFM images returned smaller particle counts per square micrometer by a factor of 5 at 1 min immersion and by a factor of 14 at 60 min immersion as shown in Fig 4.4. Watershed analysis of SEM images appears to identify small ($\leq$ 5 nm) particles more often than AFM images. Measured particle densities are most similar at 8 min, where SEM and AFM images each show monomodal distributions of 169 nm particles.

4.6 Conclusion

For proper model validation and troubleshooting, accurate image processing methods must be used to analyze SEM and AFM images. The watershed method and PHM provided a more accurate analysis of our NP samples. In addition to developing the algorithms, the comparison between the two most common acquisition methods, SEM and AFM, showed shortcomings of AFM. Result confidence was increased using automated image processing algorithms for SEM which allowed for large sample statistical analysis (over 1000 particles).

CHAPTER 5

CONCLUSION

5.1 Concept to Reality

This work has covered a wide variety of topics including NP array modeling, fabrication, and model verification through NP analysis. Some of the more noteworthy points of this work as well as novel discoveries have been outlined in bulleted form below.

- A new analytical solution to Maxwell's equation for infinite two-dimensional arrays for retarded dipole interactions between spherical nanoparticles was derived. This new solution is valid for nonmagnetic spheres that are small enough (< 100 nm) that multipole interactions can be ignored and large enough that quantum effects can also be ignored (2.29).
- The new Monte Carlo interpolation method allowed for a decreased number of function evaluations. Using this new method to replace a fixed step size in independent variables allows for larger array simulations. Using the Monte Carlo interpolation technique that was developed, $100,000,000,000$ NP 2D array simulations were demonstrated in this work and multivariable design and optimization was improved by increasing the algorithm efficiency (2.C). This new variable step size method is only accurate when the exact knowledge of the singularities are known in order to determine the proper interpolation ranges, i.e. $A = [0\sqrt{2}\pi], [\sqrt{2}\pi 2\pi]$, etc.
- An expression (2.35) capable of defining all singularities within the new analytical solution (2.29) was also obtained. The knowledge of these singularity locations and behavior offers new fundamental insight into NP array sensitivity enhancement. In addition, singularities which were previously searched for numerically can be found analytically without using costly direct search methods.
- Two-dimensional NP array interparticle spacing, NP diameter (valid for NP diameters $< 100nm$ but sufficiently large to ignore quantum effects), surrounding medium, and NP material can be tuned using the speed of the new algorithm to optimize sensor design for target wavelengths (2.8).

- In coordination with Ahn et al. [2], a new fabrication technique was developed which involves electroless deposition (EL) followed by thermal annealing to produce random gold nanoparticle arrays (3.1).
- A comparison between AFM and SEM postprocessing particle size and spacial particle density analysis was conducted. A solution was also developed to counter AFM tip-convolution using the Peak Height algorithm (Table 4.1) to produce more closely matched datasets between AFM and SEM tools.
- A particle shape analysis algorithm was invented allowing randomly oriented nanoparticles to be analyzed for shape. Particles were rotated and aligned to produce a 3D histogram of stacked particle shapes. From the 3D histogram, particle shapes could be produced that represented any desired percentile of the population allowing particle shape confidence plots.
- The watershed transform was adapted to show the particle selection benefits for touching particles compared to simpler threshold alternatives which tend to exaggerate particle size due to clumping (4.2).

Combining the new analytical solution Monte Carlo technique, EL annealing method, and analysis methods provides a total concept to reality procedure for producing optimized NP designs. NP composition, spacing, size, wavelength, and refractive index can be optimized for every scenario within the model validity range. Once the parameters have been determined for an Au array, EL annealing experimentation (Fig 4.1) can be performed. The watershed transform can then be used on SEM images to acquire particle size distributions, count densities, and shape characteristics. Automated SEMs and scatterometers used in the semiconductor industry can be used likewise to achieve real-time process control and array analysis in a manufacturing environment.

5.2 Future Work

The analytical solution is valid for a finite range of spherical NP diameters for 2D arrays due to multipole and quantum effects. Future work to improve the current model validity range by eliminating assumptions or maximize its use have been listed below.

- A general expression valid for spherical NP diameter $>100\ nm$ could be derived by replacing (2.13) with the higher multipole expression for a retarded electric field and redoing the remainder of the derivation. This new expression would then be valid for all nonmagnetic spherical 2D arrays where NP diameter size was large enough to

ignore quantum effects.

- Au and Ag are common NP array materials for sensor applications. Using the speed improvements of the new analytical solution, NP material could be optimized to search for a new theoretical material with superior optical properties. Once an ideal material was identified, the NP 2D array fabrication technique would then have to be developed as well.
- The redundancy elimination procedure for the 2D infinite NP array could be expanded to include that of a 3D infinite NP array. The derivation would be modified where the $3x3$ matrix in (2.18) would be replaced by a $3x3x3$ matrix and (2.19) by $5x5x5$ before continuing on with the derivation.
- The dilogarithm, $P_2(Z) = \sum_{n=1}^{\infty} \frac{Z^n}{n^2}$, offers a significant roadblock to the continued simplification of the analytical solution. Progress towards obtaining a dilogarithm analytical solution or symbolic regression approximation could offer additional insight and improvement.
- Adding the substrate electric field contributions to (2.13) for a more accurate solution using the principle of superposition. An other alternative is using a surrounding medium with an identical refractive index as the substrate to eliminate substrate electromagnetic interference.

Despite the benefits of the current model, future work is necessary to expand the model validity range and usefulness. With the existing model, direct search methods could be used to inversely solve the necessary material optical properties for maximum sensitivity. The desired optical properties could be compared to a database of known material, or using physical chemistry nonexistent compounds/alloys could be suggested.

REFERENCES

[1] Wikipedia, the free encyclopedia, Nov 2009.

[2] Wonmi Ahn, Benjamin Taylor, Analia G. Dall' Asen, and D. Keith Roper. Electroless gold island thin films: Photoluminescence and thermal transformation to nanoparticle ensembles. *J Opt Soc Am B*, 2:931–949, 2008.

[3] Hitachi vendor Andy. Hitachi vendor training. In *IMFT Hitachi vendor training*, July 2009.

[4] S. T. Bramwell, S. R. Giblin, S. Calder, R. Aldus, D. Prabhakaran, and T. Fennell. Measurement of the charge and current of magnetic monopoles in spin ice. *Nature*, 461:7266, 2009.

[5] Sockalingum G D, Beljebbar A, Morjani H, Angiboust J F, and Manfait M. Title not added. *Biospectrosc.*, 4:S71–S78, 1998.

[6] Spadavecchia J, Prete P, Lovergine N, Tapfer L, and Rella R. Title not added. *Phys. Chem. B*, 109:17347–17349, 2005.

[7] J Jackson. *Classical Electrodynamics*, pages 243–245. John Wiley Sons, INC., third edition, 1999.

[8] Beomseok K, Tripp S, and Wei A. Self-organization of large gold nanoparticle arrays. *J. Am. Chem. Soc.*, 123:7955–7956, 2001.

[9] Hata K, Yoshida S, Fujita M, Yasuda S, Makimura T, Murakami K, and Shigekawa H. Self-assembled monolayer as a template to deposit silicon nanoparticles fabricated by laser ablation. *J. Phys. Chem. B*, 105:10842–10846, 2001.

[10] Khriachtchev L, Heikkila L, and Kuusela T. Title not added. *Appl. Phys. Lett.*, 78:1994–1996, 2001.

[11] Shigeta M and Watanabe T. Growth mechanism of silicon-based functional nanoparticles fabricated by inductively coupled thermal plasmas. *Appl. Phys.*, 40:2407–2419, 2007.

[12] James Clerk Maxwell and Joseph John Thomson. *A Treatise on Electricity and Magnetism*, volume 2. Clarendon press, 3 edition, 1904.

[13] F Meyer. Topographic distance and watershed lines. *Sig. Proc.*, 38:113–125, 1994.

[14] Michael I. Mishchenkoa, Gorden Videenb, Victor A. Babenkoc, Nikolai G. Khlebtsovd, and Thomas Wriedte. T-matrix theory of electromagnetic scattering by particles and its applications: a comprehensive reference database. *Journal of Quantitative*

Spectroscopy Radiative Transfer, 88:357–406, 2004.

[15] Mallory G O and Hajdu J B. *Electroless Plating: Fundamentals and Applications*, chapter 1. American Electroplaters and Surface Finishers Society: Orlando, FL, 1990.

[16] Nakano S, Ishizuka M, Wang Q, and Ogiso H. Nanoparticle fabrication for micro total analysis system by using ion implantation and surface etching. *Surface and Coatings Technology*, 187:167–171, 2004.

[17] M N O Sadiku. *Elements of Electromagnetics*, page 386. Oxford University Press, fourth edition, 2007.

[18] Tapan K. Sarkar, Robert Mailloux, Arthur A. Oliner, Magdalena Salazar-Palma, and Dipak L. Sengupta. *History of Wireless*, page 22. Wiley-IEEE Press, first edition, 2006.

[19] A. Taflove, K. R. Umashankar, B. Beker, F. A. Harfoush, and K. S. Yee. Detailed fdtd analysis of electromagnetic fields penetrating narrow slots and lapped joints in thick conducting screens. *Antennas and Propagation, IEEE Transactions*, 32:247–257, 1988.

[20] Ben Taylor, Braden Harbin, and Keith Roper. Exact solution to the coupled dipole approximation of maxwell's equation for electromagnetic interactions in periodic nanostructures. *PNAS*, Submitted 2011.

[21] Maxim A. Yurkin and Alfons G. Hoekstra. The discrete dipole approximation: an overview and recent developments. *J.Quant.Spectrosc.Radiat.Transf.*, 106:558–589, 2007.

[22] Schatz GC Zou S. Theoretical studies of plasmon resonances in one-dimensional nanoparticle chains: narrow lineshapes with tunable widths. *Nanotechnology*, 17:2813–2820, 2006.